Recovery of the geometric road design Elements using low-cost sensors: a feasibility study

Von der Fakultät für
Luft- und Raumfahrttechnik und Geodäsie der Universität Stuttgart
zur Erlangung der Würde eines Doktor-Ingenieurs (Dr.-Ing.)
genehmigte Abhandlung

Vorgelegt von
Godfrey Onyango Ogonda, M.Sc.
aus Kisumu, Kenia

Hauptberichter:	Prof. Dr.-Ing. Alfred Kleusberg
Mitberichter:	Prof. Dr.-Ing. Christian Lippold
Tag der mündlichen Prüfung:	9. März 2009

Institut für Navigation der Universität Stuttgart

2009

ABSTRACT

Roadway inventory databases are continuously being applied in two main areas, first in the design, construction and maintenance of new and existing roads and secondly in the specific areas of intelligent transport systems such as driver assistance and mobile information systems. The initial development, storage and management of the database have therefore become the responsibility of the various road agencies around the world, which ascertain that this is done in accordance with the respective road policies and design standards in the respective states or regions. The roadway inventory database has been estimated to comprise of about 49 possible collectible categories of roadway inventory data elements such as the horizontal curve elements, the vertical alignment elements, drainage, the number of lanes and the traffic signs among others. As a result of the rapidly changing roadway environment, the challenge therefore is the up-dating of the roadway inventory database using relatively accurate and cost-effective techniques, as opposed to the mobile mapping systems usually employed in their initial development. This calls for investigation into other low-cost data collection and processing methods and devices. The aim of this study is therefore to develop a methodology for extracting the geometric road design elements from position time series derived from a set of low-cost sensors, and to assess the accuracy potential of such sensors.

The report begins with a review of the basics of road design and the formulations used in the estimation of the geometric road design elements as well as a brief mention of the road design standards and guidelines. The various sensors and systems for land based positioning and navigation, with emphasis on their costs and sources of error, are discussed. An overview of the theory of curve fitting functions and criteria with specific interest in splines and their properties is also presented.

The investigations, tests and data analyses carried out are based on computer simulated data as well as on field measured data collected in the vicinity of the city of Stuttgart. A set of spline-based algorithms is developed and tested for the extraction of geometric road design elements for the purpose of updating the roadway inventory database. Furthermore, a set of low-cost navigation sensors, comprising of wheel speed sensors, a single-axis micro electro-mechanical systems accelerometer and a single-axis micro electro-mechanical systems gyroscope, was built up and utilised for the acquisition of vehicle track positions. The methodologies and formulations used in the synchronisation of the measurements and the calibration of the sensors, the integration of the sensors' data based on a Kalman filter algorithm, the estimation of curvature and recovery of the geometric road design elements are presented.

In summary, the tests and data analyses illustrate that

- the developed spline-based algorithms successfully extract the geometric road design elements based on the assumption that the road geometry consist of straight lines, circular and transitional curves, as demonstrated by the analysis of simulated data.
- the field calibration procedure of the wheel speed sensors employing a Kalman filtering process achieve an accuracy of $\pm 0.04\ m$ in distance increments. The derived heading differences have a final cumulative deviation of $\pm 3°$ or a drift of about $0.0025\ °/s$.
- the proposed set of low-cost sensors can be a viable solution as a short duration georeferencing device for mobile mapping systems in the recovery of the geometric shape of the road track, as long as (i) the vehicle is driven at low speed (about *36km/h or 10m/s*) but certainly above the threshold speed of the wheel sensors, (ii) the random

errors in the measurements are characterized and reduced and (iii) the pitch angle and gravitation effect are compensated for in the measured data.

- the random and the systematic errors in the measured vehicle track position time series are a serious predicament to curvature estimation, which however can be reduced by the use of smoothing splines and by calibration of the measuring devices, respectively.

Based on the above tests and analyses, the low-cost sensors together with the developed algorithms generated sufficiently accurate geometric road design elements for updating roadway inventory databases especially those intended for intelligent transport system applications whose accuracies are specified in metres to sub-meter. However, the recovered geometric road design elements may not be sufficiently accurate for roadway inventory databases used for road design and construction because here the accuracy requirements are in centimetre level.

Zusammenfassung

Die erfassten Streckendatenbanken werden laufend in zwei Hauptbereichen angewandt, erstens bei der Ausführung, Bauweise und Instandhaltung von neuen als auch der schon bestehenden Strecken und zweitens in bestimmten Bereichen intelligenter Transportsysteme, wie z. B. die mobile Informationssysteme. Infolgedessen tragen die verschiedenen Organisationen für Straßen und Verkehrswesen rund um die Welt die Verantwortung für die Anfangsentwicklung, die Archivierung und den Betrieb der Datenbanken und dafür, dass dies in Einklang mit der jeweiligen Straßenbaupolitik und dem Ausführungsstandard des jeweiligen Landes oder der jeweiligen Region bewerkstelligt wird. Die erfassten Streckendatenbanken enthalten schätzungsweise 49 mögliche erfassbare Streckendatenkategorien, wie z. B. horizontale Kurvenelemente, vertikale Anpassungselemente, Abflusskanäle, die Anzahl der Straßen und Straßenschilder. Die schnell wechselnde Streckenumgebung stellt deswegen eine Herausforderung für die Aktualisierung der erfassten Steckendatenbanken dar, wobei relativ genaue und kostengünstige Verfahren verwendet werden im Gegensatz zu dem kinematischen Vermessungssystem, das ursprünglich für die Entwicklung der Streckendatenbanken verwendet wurde. Dies fordert eine Untersuchung von anderen preiswerten Datenerfassungs- und Verarbeitungsmethoden sowie der jeweiligen Geräte. Die Zielsetzung dieser Untersuchung ist infolgedessen die Entwicklung einer Methode zur Extrahierung geometrischer Straßenbauelemente aus einer Zeitreihe von Positionen, die von einem Satz preiswerter Sensoren abgeleitet werded und die Einschätzung des Genauigkeitspotentials solcher Sensoren.

Der Bericht beginnt mit einem Überblick über die Grundlagen der Streckenausführung und den Formulierungen, die im Schätzverfahren der geometrischen Straßenbauelemente verwendet wurden und einer kurzen Übersicht der Straßenausführungsstandards und –richtlinien. Weiterhin werden noch die verschiedenen Sensoren und bodengestützten Positionierungs- und Navigationssysteme erwähnt, wobei auf deren Kosten und Fehlerquellen besonderen Wert gelegt wird. Außerdem enthält der Bericht eine Übersicht über die Theorie der Kurvenanpassungsfunktionen und –kriterien, mit einem besonderen Augenmerk auf Splines und ihre Eigenschaften.

Die ausgeführten Untersuchungen, Tests und Datenanalysen stützen sich sowohl auf computersimulierte Daten als auch auf Feldmessdaten, die in der Umgebung der Stadt Stuttgart erfasst wurden. Ein Satz splinebasierter Algorithmen zur Extrahierung geometrischer Straßenbauelemente wurde entwickelt und getestet, um die erfassten Streckendatenbanken zu aktualisieren. Des Weiteren wurde ein Satz den preisgünstigen Navigationssensoren gebaut, das aus Radsensoren, einem einachsigen mikro-elektro-mechanischen Systeme (MEMS) Beschleunigungsmesser und einem einachsigen mikro-elektro-mechanischen Systeme (MEMS) Gyroskop zusammengesetzt ist. Dieser Prototyp wurde zur Akquisition von Fahrzeugwegpositionen verwendet. Es wird des Weiteren folgendes aufgezeigt: die Methodik und die Formulierungen, die bei der Synchronisation der Messungen und der Kalibrierung der Sensoren verwendet wurden, die Integration der Daten mittels der Kalman-Filter-Algorithmen und die Bewertung von Krümmungen bzw. die Extrahierung geometrischer Straßenbauelemente.

Kurz gefasst zeigen die Tests und Datenanalysen folgendes:

- die entwickelten splinebasierten Algorithmen extrahieren erfolgreich die geometrischen Straßenbauelemente ausgehend von der Annahme, dass die

Streckengeometrie aus Gerade, Kreis- und Übergangsbogen besteht, wie aus der Analyse der simulierten Daten hervorgeht.

- das Feldkalibrierungsverfahren der Radsensoren, die ein Kalman Filter Verfahren verwenden, erzielen eine Entfernungsinkrementgenauigkeit von $\pm 0.04\ m$. Die abgeleitete Richtungsänderung hat eine endgültige kumulative Abweichung von $\pm 3°$ oder eine Richtungsverschiebung von etwa $0.0025\ °/s$.
- die vorgeschlagenen preisgünstigen Sensoren sind eine praktikable Lösung für ein Georeferenzierungsgerät mit ausreichender Relatiogenauigkeit für mobile Meßsysteme in der Rückgewinnung der geometrischen Form der Strecke, solange das Fahrzeug sich mit einer niedrigen Geschwindigkeit (etwa *36km/h oder 10m/s),* jedoch über der Grenzgeschwindigkeit der Radsensoren bewegt; (ii) die statistischer Fehler der Messungen werden charakterisiert und verringert (iii) wobei der Nickwinkel und der Gravitationseffekt in den gemessenen Daten berücksichtigt werden.
- die zufälligen und die systematischen Fehler der gemessenen Zeitreihe von Positionen sind ein Hindernis in der Bestimmung der Krümmungen, das jedoch durch die Verwendung von glättenden Splines und durch Kalibrierung der Messgeräte umgangen werden kann.

Durch die oben genannten Tests und Analysen wird gezeigt, dass die preisgünstigen Sensoren mit den entwickelten Algorithmen genügend genaue geometrische Straßenbauelemente zur Aktualisierung der erfassten Streckendatenbanken, insbesondere solche für intelligente Transportsysteme, deren Abweichungen in Metern gemessen werden, erstellt haben. Trotzdem könnten die erstellten geometrischen Straßenbauelemente nicht genau genug sein für die in der Straßenausführung und dem Straßenbau erfassten Streckendatenbanken, weil hier die Genauigkeit im Zentimeterbereich liegt.

TABLE OF CONTENT

LIST OF FIGURES

LIST OF TABLES

1 Introduction

Equation Section 1

A navigation solution involves the estimation of the present position, velocity and orientation (attitude) of the vehicle sometimes with the aim of reaching a waypoint in the shortest time or shortest distance safely. Although the accuracies, $\pm 13m$ in position at *95%* confidence interval (US DoD, 2001), achieved by the global positioning system (GPS) standard positioning service and the bounded errors provided in position and velocities are sufficient for the required navigation performance (accuracy, reliability, continuity and availability), most land based navigation applications require better positional accuracy than can be provided by a single GPS receiver. To achieve the desired accuracy, an augmentation technique commonly known as the differential global positioning system (DGPS) is usually applied. The DGPS technique is based on applying corrections, as obtained from a highly accurate geodetically surveyed location of a GPS reference station, to the roving receiver. DGPS technique is still subject to multipath, signal blockage or obstruction and interference especially in urban environment, under bridges, in tunnels or forests. These usually result in fewer than four satellites, poorly elevated satellites, poor signal to noise ratio or poor geometric constellation. Therefore, the DGPS also falls short of the required navigation performance. The preferred choice in alleviating temporarily degraded DGPS position loss is usually integrated navigation system, which is a combination of the on board navigation solution as derived from inertial sensors such as accelerometers, gyroscopes etc. and other independent navigation data (navigation aides) to update or correct the navigation solution. The inertial sensors provide the navigation solution while the global navigation satellite system such as GPS is commonly used as aides to navigation. The use of integrated navigation systems in kinematic surveys of the vehicle track and road geometry measurement offers two main advantages. First, the signal outages, obstruction or interference degrades the positions derived from the DGPS. Secondly, the DGPS bandwidth for position output is limited with most receivers giving position sample at a rate of *1Hz*. Mounted on a measuring car travelling at *70 km/h* i.e. about *19 m/s*, the receiver will sample position every *19 m*. These data sampling rate is too low for accurate recovery of geometric parameters.

Currently land navigation systems seek to improve efficiency (time and cost), mobility and safety by providing users with the position, velocity and other geographical information in real time. The information provided by the navigation systems are stored in digital geographical information system (GIS) databases as attributes associated to features (e.g. roads, services or administrative boundaries) and relations (e.g. turn restrictions). The advanced utilization of the information technology within the road transport field, has resulted in the modern navigation systems providing several other services, a concept referred to as intelligent transport systems (ITS). The two types of ITS are the *driver assistance (DA) systems* (e.g. cruise controls, lane departure warning systems, obstacle and curve warning systems) aimed at enhancing road safety and the *Mobile Service and Information (MSI) systems* (navigation systems, mayday and emergency assistance, electronic fee payment, real-time traffic information and driver specific convenience adjustments) aimed at making travelling more pleasurable, productive and efficient. The advanced driver information and warning systems require (i) detailed and accurate road geometry data to be included in the databases for example through creation of a 3D digital terrain model obtained by georeferencing of photographic images and laser scanning (ii) the provision of the

information in real time for example through integrated navigation systems and (iii) the dispensation of the roadside warnings and information to the users.

Integrated navigation systems combined with information technology (IT) are products of modern communication, computer networks, GPS and must therefore maintain high precision and low redundancy database. Hatger and Brenner (2003) stated that, today's car navigation systems have reached a high level of maturity requiring huge map databases with wide coverage and up-to-date. These applications require that the information on the true road geometry to be added onto the database. Apart from being up-to-date and accurate, the information should also be in usable format and contain other details such as addresses, hotels, parking, filling stations, restaurants etc. The major components of such roadway databases are road networks, towns and villages, public transport and topographical maps (details). Typically road network data contained in the roadway database is an integrated and representative database in a digital environment of the geometry and thematic characteristics of the road axis.

In the geometric design and construction of roads the design standards and regulations issued by state agencies and research organisations must be followed. For example, in Germany several documents spell out the design procedures to be followed, examples of which include *Richtlinien für die Anlage von Straßen (RAS)*: Teil *Linienführung (RAS-L)*, *Empfehlungen für die Anlage von Hauptverkehrsstraßen (EAHV 93)*, *Richtlinien für die Anlage von Straßen (RAS)*, Teil: *Querschnitte (RAS-Q)*, for the horizontal, network and cross section designs respectively. The implementation of the design standards are not uniform even on a national basis due to variations in laws and financial capabilities and other policy issues of the different governmental agencies. Other regional factors such as terrain and weather conditions also have an influence on the implementation phase. An ideally designed road should be in harmony with the existing natural topography so as to cut on the construction costs. This is usually not possible due to the different design aspects which must be taken into consideration. But an accurate and timely road inventory data is capable of supporting the planning, design, construction and maintenance of a variety of the other transport facilities. Thus, to improve on the management of roads, it is necessary to integrate the database and keep it up to date. Technological developments in the navigation systems, GIS and ITS are also currently influencing geometric road design practices in many countries where for example reliance on traffic road signs to convey information is being faced out by adoption of design policies and procedures which enhance self-explaining, self-enforcing roads. Roads are being designed for specific purposes within the context of natural and human environment by employing aesthetics to explain the road functions and enforce speed for efficient mobility and safety among other factors. This is a new concept, i.e. tailoring the road design to local project circumstances, is referred to as context-sensitive design (CSD); a term coined in the United States of America and currently being applied by many road design agencies. The CSD concept requires a detailed roadway inventory database which includes various traffic signs, drainage patterns, horizontal curve elements, vertical alignment, lanes, sideways, bridges and several other details as listed by Khattak *et al*. (2001). The geometric model consists of linear segments, circular arcs and spiral curve (clothoids) while the thematic characteristics include the qualitative and quantitative attributes such as signs, elevations, slopes and radii of curvatures.

Several factors must therefore be considered in the development, storage and management of the roadway inventory database to enable their useful application in the design, construction and ITS. The accuracy, for example, of both the georeference and descriptive information

determines the usefulness of the roadway inventory database to a significant extent. The accuracy of the descriptive inventory data has been shown to depend on the method of collection. Until recently, compilation of roadway database was done by traditional manual methods, typically involving a hand held GPS and conventional land surveying techniques, although providing higher positional accuracy, are time consuming and labour intensive. Due to the rapid evolution in the roadway environment, the roadway data collection is currently being achieved through improved information collection and provision using such systems as mobile mapping systems. A mobile mapping system (MMS) uses a set of georeferencing and imaging sensors combined with software techniques to measure the position of objects and capture images along the roadway. A number of research works have shown that the accuracy of inventory elements collected by different MMSs are sufficiently high for roadway inventory database (Coetsee *et al.*, 1994; Vaidya *et al.*, 1994; Schwarz and El-Sheimy, 1999), and consequently several roadway agencies using the manual methods of data collection are considering adopting the mobile mapping systems.

1.1 Statement of the problem

Psarianos *et al.* (2001) observed that while sufficient information on geometric design of newly constructed roads may be available, there tend to be considerable lack of data regarding older roads constructed and maintained under various road jurisdictions. Since road networks play a vital role in the economic development of a country, maintaining high precision, up-to-date and low redundancy roadway inventory database of the road network, their geometrical properties and other details within the roadway environment such as hospitals, gas stations etc. is therefore a critical national infrastructure activity of many road transport agencies around the world. The road transport agencies develop, maintain and regularly update vast inventories of a variety of roadway elements and have different requirements for the inventory database or ITS. The challenge therefore is to find accurate and low-cost data collection and processing methods and devices to update the different roadway inventory database elements. This is done in accordance with the respective road policies and design standards in the states or regions. Khattak *et al.* (2001) identified *49* possible roadway inventory elements from *24* possible data categories that are collectable by a MMS employing digital image capture techniques. These included the horizontal curve category (radius and curve length), the vertical alignment category (tangent grades and lengths), drainage, lanes and traffic signs among others. Technological viability, accuracy and cost effectiveness of the data collection technologies is a priority in the data collection and development of the roadway inventory database. This calls for further research into the accuracy, reliability and cost-effectiveness of the data collection methods and devices. Current techniques attempt to extract the road network and road geometry from the laser scanned images obtained by the expensive more accurate MMS which integrate multi-sensor systems and high performance computers. The first main objective of this research was therefore **to develop a methodology for extracting the geometric road design elements from position time series** for the purpose of updating or incorporation of the geometric road elements into the roadway inventory database.

The high position accuracy required in roadway inventory databases cannot be provided by GPS alone as the georeferencing technique for the MMS. The mobile mapping and navigation systems must provide accurate position solutions even during GPS dropouts. The most accurate and preferred approaches to aiding GPS navigation are either by map matching or integration with dead reckoning (DR) sensors. Integrated navigation systems are therefore preferred due to the complimentary nature of the sensors involved in the system. As components of the integrated navigation systems, the inertial sensors are subject to low

frequency errors while the GPS is subject to high frequency errors and their integration is performed using the Kalman filter algorithms to supplement each other with additional position accuracy. One major factor that has limited the application of the more precise inertial sensors in the land based integrated navigation systems is their cost. The disadvantage of the high costs of inertial navigation systems have been surpassed by the development of low-cost micro-electromechanical (MEMS) sensors. The improvements in low-cost MEMS sensors with relatively better stability and accuracy and the fact that anti-lock breaking systems (ABS) are almost standard in many vehicles today, it is hoped that their incorporation into the navigation system can provide an opportunity to tremendously cut down the costs of updating the geometric road design elements contained in the roadway inventory database by acquisition of the actual geometry of the road axis after construction or during maintenance. Although the accuracies of the measured geometric track data depend on the accuracy of the integrated navigation system being used in the measuring process, our proposition is to use low-cost integrated navigation system. The position data obtained using such low-cost systems contain noise and highly correlated errors. The established concepts or methods used in digital signal processing, optimal estimation and data reduction are applicable in the estimation of the road geometry from measured data points but the error distribution and noise structure in the measured data points obtained by the low-cost navigation systems doesn't allow their direct application. On this background, the second main objective of this research was **to investigate the possibilities and limitations of using low-cost sensors integrated with GPS to determine position time series of sufficiently accuracy for the above purposes**. These are considered on the basis of achieving the accuracy acceptable to the roadway inventory data base, their reliability and cost effectiveness. Some of the low-cost sensors considered in these investigations were

- MEMS gyroscope
- MEMS accelerometer
- anti-lock breaking system (ABS) wheel sensors
- GPS

This thesis presents a prototype low-cost integrated navigation system for obtaining positions, especially of the horizontal profile of the vehicle, together with a set of algorithms for data filtering and robust extraction of the geometric road design elements for subsequent transfer to the road inventory database. The analysis and findings based on simulated data, field measurements collected around Stuttgart city area and the developed algorithms are also presented.

1.2 Organisation of the thesis

The thesis is organised into seven chapter and two appendices. Chapter one is a brief introduction of the general theme of the thesis and outlines the research objectives. An outline of the organisation of the thesis is also described in this chapter.

Chapter two discusses the basics of road design and the formulations used in the estimation of the geometric road design elements. The chapter also reviews the different road classes as covered by the design standards and guidelines outlined by the German roads agencies.

Chapter three highlights the various sensors and systems applicable in land based positioning and navigation with sections devoted to inertial navigation systems, anti-lock braking system wheel speed sensors and mobile mapping systems. Their varied applications and sources of error are also covered.

Chapter four gives an overview of the theory of curve fitting functions and criteria. Discussion of the different classes of splines and their properties is also presented in this chapter. The concepts of curve differential geometry and curve parameterization are briefly reviewed.

Chapter five presents the results of the investigations carried out based on the simulated data to determine the effects of random noise and systematic errors in sensors on the recovery process of the geometric road design elements. In the chapter, the findings of the tests on the proposed geometric road design elements recovery algorithm are also presented.

Chapter six is a continuation of the testing of the geometric road design elements recovery algorithm and presents the proposed low-cost integrated navigation system. The chapter also describes the synchronisation of measurements and calibration of the sensors used in the collection of field data. The integration of the measurements based on a Kalman filter algorithm is also described. The results of recovery of the geometric road design elements from the field data collected around Stuttgart city are also contained in this chapter.

The summary and conclusions drawn from the analysis and findings are contained in chapter seven.

2 Geometric Road Design Elements

Equation Section 2

The objective of this study was to the develop a methodology for describing an existing road section in the most compact way using the geometric road design parameters/elements derived from a low-cost integrated navigation system. The basics of road design and construction, presented briefly in this section, was therefore a prerequisite.

2.1 The basics of geometric road design

Design of roads and other transport systems involves three main steps. The first part involves the initial physical planning aspects of the proposed system and is basically concerned with the determination of need, location and time of the intended project. The second part involves the design which basically deals with the precise location of the facility, size and level of performance to be achieved. The design standards are categorised into two; *geometric* and *structural* design standards. The geometric design aims at achieving maximum safety by optimizing cost, topography, design speed and traffic capacity and consists of cross section, vertical- and horizontal alignments. The primary concern of the geometric design is to relate the physical elements of the highways to the requirements of the driver and the vehicle. The geometric design parameters are also used in the construction of a new road, alterations and/or extensions of existing roads to harmonise the road to the surrounding topography and to fit the road to the economical and physical environmental factors in the location. The structural design standard, on the other hand, consists of pavement (load bearing and drainage) design and is thus concerned with the thickness of the flexible or rigid pavement. The third part in the design process is the monitoring of the facility through costing, use optimisation and regulations.

Different factors influence the elements of design and construction of roads. These factors can be classified into three broad categories; (i) environmental factors i.e. the physical, biological and aesthetics aspects of the surrounding, (ii) the economic factors i.e. funding, public comments, land use and security (iii) the traffic quality of the road i.e. function, average daily traffic (AVT), design hourly volume (DHV) and design speed. These factors are not independent of one another and must be optimised in the design process although more weight is put on the traffic quality since it is the main road classification factor. The basic objective of the design process is to create a road with proper dimensional values and alignment characteristics by balancing the environmental and economical factors so as to achieve the desired traffic quality at peak hours without extravagance or wastage.

The road can be described geometrically in the longitudinal and traversal by four main elements. These are (i) the straight lines/portions (ii) both horizontal and vertical circular curves (iii) non-circular transition curves and (iv) the cross section elements. The elements have to be optimised in such a manner such that there is small or no position, direction and curvature discontinuities between the different geometric entities (Hatger and Brenner, 2003). The elements are limited to certain ranges depending on the design speed. Apart from the geometric road elements, other attributes along the road corridor are equally important. They can be used to inform, guide and warn the driver of the circumstances surrounding the route of travel. Detailed location of physical features abutting the roadway including the drainage

structures, roadway signs, lighting facilities and right of way boundary marks must also be located and coded into the data collection system.

Design and construction of roads is usually done using laid down standard methods and techniques and these design standards are contained road design guidelines for different countries and regions world over. In Germany, for example, the road design standards are contained in a publication called *Richtlinien für die Anlage von Straßen (RAS)*. A sample case of the German road design and construction guidelines are shown in Table (2.1). In this case, the basic principles, methods as well as limits and standard values of the road design elements for the new, alteration and extension of roads in rural and urban areas are described in detail by several documents, viz. RAS-L (1995), RAS-Q (1996), RAS-N (1988), RAS-K-1 (1988), EAHV 93 and EAE 85/95, depending on the category (class) of the road being designed. RAS-L(1995) sets guideline for the project planning and design stages, design speeds, horizontal and vertical alignments, cross sectional design and sight distance design elements. RAS-Q(1996) spells basis for the assessment of the constituent elements of road cross sections, road cross section layout and verification of traffic quality. RAS-L (1995) and RAS-Q(1996) are applicable in road classes *AI-VI* and *BI-II*. EAHV (1993) and EAE (1985,1995) contain recommendations for the design, construction and layout of collector roads and motorways for road classes *BIII, CI-III, DIII* and *BIV, CIV, DIV*, *EIV* respectively (*cf*. Table 2.1). RAS-N (1988) specifies the concepts of road network designs while RAS-K-1 (1988) state the principal solutions and design basics for road junctions.

Table 2.1: *Applicable areas of RAS-L, EAHV and EAE for road classes in RAS-N (After Ressel, 1999)*

Road Category / Road Function Level		Outside built-up areas	Within built-up area			
		Rural motorways	Urban roads			
		Mobility			Access	Public Realm
		A	B	C	D	E
Extensive road networks	I	AI	BI	CI	-	-
Regional road networks	II	AII	BII	CII	-	-
Inter-municipality roads	III	AIII	BIII	CIII	DIII	-
Expansive closed-loop roads	IV	AIV	BIV	CIV	DIV	EIV
Non-systematic roads	V	AV	-	-	DV	EV
Route connections	VI	AVI	-	-	-	EVI

RAS-L, RAS-Q | EAE 1985/95 | EAHV 1993 | - None is Applicable

2.2 Formulation of geometric road design parameters

The geometric road design elements are sight distances (passing sight distance, stopping sight distance), horizontal alignment (superelevation and curvature) and vertical alignment (cross sections, separate grade lines and vertical curves). Thus the geometric location of a road can

be quantified by bearing, distances, curvature, transitions, stations and offsets, all of which are expressed in relation to the road centre line. The physical elements that must be taken into account in geometric design are

- Horizontal radius of curvature
- The length and parameter of the curves
- The starting and ending point of the various components
- Superelevation
- Vertical radius of curvature
- Grades

The horizontal and vertical alignments discussed briefly in this section are based on the assumption that there are no discontinuity in curvature i.e. no abrupt reversal of curves. As listed by American Association of State Highways and Transport Officials (AASHTO) Green Book standards, compound curves have also been avoided because they are rarely used in road design.

The design of the individual physical element is influenced by the design controls which include characteristics of the vehicles, traffic and drivers. These elements have been quantified and justified through design standards set out by respective countries' or regions' road authorities so as to achieve safety and efficiency. The design standards are further classified depending on type of road (whether dual or single carriageways roads), terrain pattern and location (urban or rural environment). The tolerance specifications for the road alignment parameter are in metres unlike railroad in which they are in tens of a millimetre. Although the geometry of the road is three dimensional, the horizontal and vertical components of road alignment are treated separately and subsequently combined and optimized at the implementation stage.

The *design speed* is the maximum safe speed that can be maintained over a specified road section when the conditions are so favourable that the design features of the highway are applicable. The design speed is a determinant factor of several horizontal and vertical geometric road design elements as well as the minimum radius of curvature, the clothoid parameter and the shape of crests and sags in vertical curves. Design speed is therefore a principal factor in the control of horizontal alignment and should ideally be measured as the 85^{th} percentile speed of the traffic approaching the geometric element.

2.2.1 Horizontal alignment elements

The horizontal alignment is performed in order to achieve a perfect relationship between the design speed, curvature and superelevation. The 2-dimensional geometry of the road is defined by three main elements of the horizontal alignment consisting of a sequence of tangents or straight lines, circular curves connected by non-circular or transitional curves (Figure 2.1). Other factors affecting selection of horizontal design elements are topography, land use, cost, safety (sight distance) and road class (design speed).

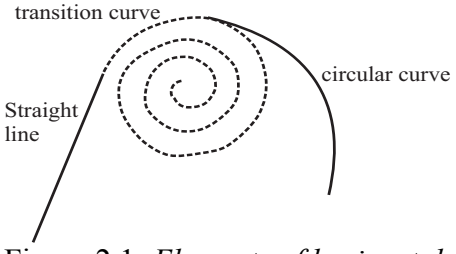

Figure 2.1: *Elements of horizontal alignment*

2.2.1.1 Tangents

For a very long time, straight lines were the most preferred road design element. Long straight-line portions of the road in the road class A were realised to have the following disadvantages:

- Difficulty in estimating the distance and speed of the approaching as well as the following vehicles
- Serious blinding by the approaching vehicles in the night
- Difficulty in fitting the road to the physical environment (terrain)
- Easy acquisition of excessive speeds by the drivers

Due to these reasons, it is recommended that the long stretches of road in category A should be minimised and replaced by curves of large radii. Due to blinding in the night, it is stipulated that long straight portion between two successive curves should be avoided. In cases where this is not possible, the length of the straight portions should not exceed 20 times the design speed, v (AASHTO, 1994).

Track bed roll angle is the difference in elevation of the road edges, a quantity referred to as the *superelevation*. The straight lines or tangents characterized by zero curvature and the track bed roll angle is typically zero i.e. no superelevation.

2.2.1.2 Circular curves

Circular curves have constant radius of curvature, constant track bed roll angle and constant non-zero curvature. The following formulations have been adapted from FGSV (1995), Muskett (1995) and Wright (1996). The degree of curvature of the curve $D\left[^{\circ}\right]$ is defined as the angle subtended at the centre by a *30.48 m (100ft)* long arc.

$$D = \frac{30.48}{R} \cdot \frac{180}{\pi} = \frac{1746.375}{R} \tag{2.1}$$

Elements of circular curve design include point of intersection of tangents, deflection angle, tangential points, radii of curvature and centre point of the curve (*cf.* Figure 2.2). Once the radius of the curve R and the angle subtended at the centre θ is also known, all the other parameters of the curve can be derived. Several variations of the circular curves are employed and include simple, compound and reverse curves.

The radius of the curve should be chosen such that a drive at the design speed over the length of the curve should last more than two seconds, the agreed average reaction time of the driver (AASHTO, 1994). The radius of the curve is also dependent on the design speed, the maximum superelevation rate and the side friction factor. The expression for minimum radius $\left[m\right]$ of the curve is given in Equation (2.2) where v is the design speed in *km/h*, e is the rate of superelevation or crossfall in percentage and f is the coefficient of friction typically varying between *0.15 (v =120 km/h*) to *0.33 (v =30 km/h*) (After FGSV, 1995).

$$R = \frac{v^2}{127\left(\frac{e}{100} + f\right)} \tag{2.2}$$

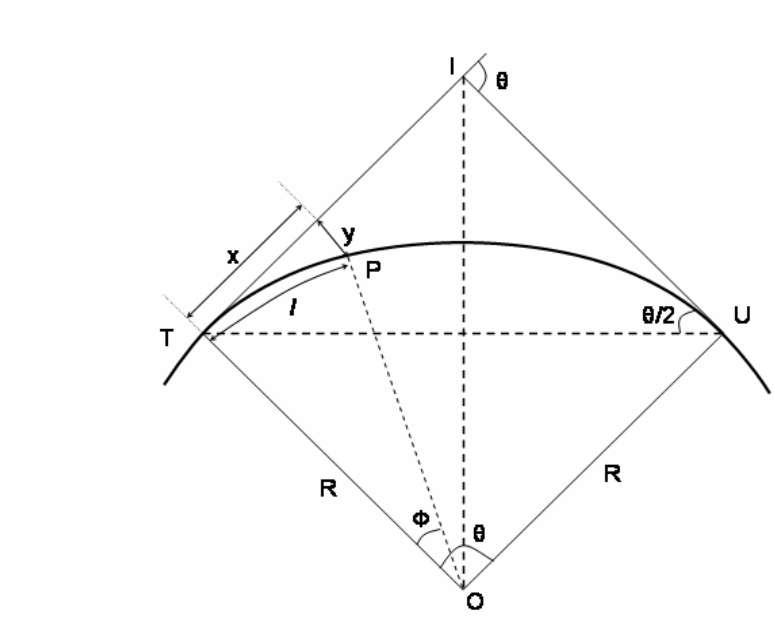

Figure 2.2*: The horizontal geometry of a simple curve (After Muskett, 1995)*

Figure (2.2) shows the geometry of a simple horizontal curve used to connect two straight lines where

θ is the total deflection angle in degrees
ϕ is the deflection angle to any point on the curve in degrees
R is the radius of curvature
T, U are the points of tangency (start and end of curves)
I is the point of intersection of tangents
P is any point on the curve
x is the tangent distance from T to P
y is the tangent offset for P

x and y therefore describe a local coordinate system with the origin being the start of the curve. The total length of the curve $L = \overline{TU}$ and the tangent length $\overline{TI}$ are computed as

$$L = \frac{2\pi R\theta}{360} \tag{2.3}$$

$$\overline{TI} = R\tan\frac{\theta}{2} \tag{2.4}$$

the length of the curve from the tangent point T to any point on the curve l, the corresponding tangent distance x and tangent offset y are given by

$$l = R\phi\frac{2\pi}{360^\circ} \tag{2.5}$$

$$x = R\sin\phi \tag{2.6}$$

$$y = R(1-\cos\phi) \tag{2.7}$$

The circular curve model in Equation (2.2) above is used to estimate minimum curve radii and minimum curve lengths in relation to design speed v and a constant superelevation of *6%* and a friction factor of *0.25* for roads of category *A* and *BII* as shown in Table (2.2). It is worth mentioning that due to lengthy calculations, most computations carried out in road design are often based on such tables that give the corresponding pre-calculated elements of the curve based on some agreed design parameters, say minimum radius and minimum curve length corresponding to a design speed.

Table 2.2*: Minimum Radii and Lengths of Circular Curves (Ressel, 1999)*

V_e	min R (m)	min L (m)
50	80	30
60	120	35
70	160	40
80	250	45
90	340	50
100	450	55
120	720	65

2.2.1.3 Transitional Curves

Transition curves are used in the horizontal design of routes mainly to accomplish the gradual transition of curvature and superelevation from the straight to the circular curve. As the vehicle path starts to curve, the lateral force, expressed in Equation (2.8), increases instantaneously from *0* to f resulting into a jolt. This can be dangerous especially if R is small and f is high. To reduce this effect, non-circular transitional curves are adopted. Therefore, the main advantages of transitional curves in horizontal curve alignment are that they provide easy flow, superelevation transition and smoother appearance of the curves (Brockenbrough and Boedecker, 1996).

Centripetal (radial) force, f, acting on a vehicle travelling at a constant speed on a circular curve is given by the following expression

$$f = \frac{mv^2}{R} \tag{2.8}$$

where

m is the mass of the vehicle
v is the velocity of the vehicle
R is the radius of the curved route

Clothoid

As mentioned above, the function of the transitional curve is to introduce the lateral accelerations gradually say from *0* to v^2/R in a distance L i.e. the total length of the transitional curve. The time taken to travel the curve length, assuming uniform velocity is $t = L/v$. Assuming a constant rate of change of acceleration q then

$$q = \frac{v^2/R}{t} = \frac{v^2}{R} \cdot \frac{v}{L} = \frac{v^3}{RL} \tag{2.9}$$

or

$$L = \frac{v^3}{Rq} \tag{2.10}$$

The minimum length of the transition curve adopted in route design is therefore determined using the rate of change of superelevation being introduced or the acceptable change of centrifugal acceleration. The common types of non-circular curves used in horizontal alignments are the lemniscates curve and the spirals examples of which include the clothoid (cornu or Euler's spiral), the cubic spiral and the equiangular (logarithmic or Bernoulli's) spiral (Kobryn, 1991). Due to computational complication, the cubic spiral, cubic parabola and Bernoulli's lemniscate were earlier used as approximation of the transition curve. With the current advancements in electronic calculations, this is no longer the case (Muskett, 1995).

The clothoid (*cf.* Figure 2.3), also known as cornu or Euler's spiral, is defined by the fundamental principle that its radius of curvature at any given point varies inversely as the arc-length measure from the beginning of the spiral. The clothoid is therefore considered as the ideal transition curve and is a preferred choice in the design of horizontal curve alignment of roads. If L is the total length of the clothoid, l the length from the start to any point on the clothoid, R the radius at the end of the clothoid and r the instantaneous radius at any point on the clothoid, then the fundamental principle of the clothoid is as expressed in Equation (2.11).

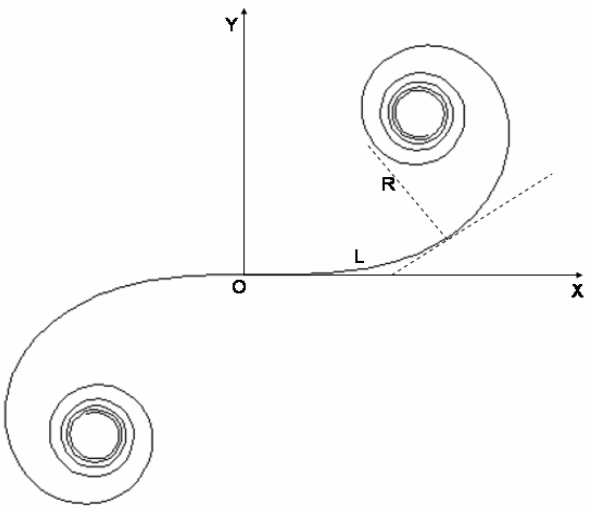

Figure 2.3: *Clothoid*

$$lr = LR = A^2 = \text{Constant} \tag{2.11}$$

where A is the clothoid parameter or the flatness of the clothoid. The length of the spiral from tangent to any point, l may be a value between zero and the total length of the transitional curve, L. The spiral deflection angle ϕ is defined as the angle between the tangent T and the tangent to the spiral curve at any point. The instantaneous curvature of the clothoid can therefore be estimated using Equation (2.12).

$$\kappa = \frac{d\phi}{dl} = \frac{1}{r} = \frac{l}{LR} \tag{2.12}$$

Integrating the spiral deflection angle with respect to the curve length one obtains Equation (2.13) below.

$$\phi = \int d\phi = \int \frac{l}{LR} dl = \frac{l^2}{2RL} + C \tag{2.13}$$

When the spiral length $l = 0$, $\phi = 0$ which means $C = 0$. The intrinsic equation of the ideal transitional curve is therefore given by

$$\phi = \frac{l^2}{2RL} \tag{2.14}$$

The rate of change of curvature of the clothoid with respect to the spiral length is a constant given by

$$\frac{d^2\phi}{dl^2} = \frac{1}{RL} = \text{constant} \tag{2.15}$$

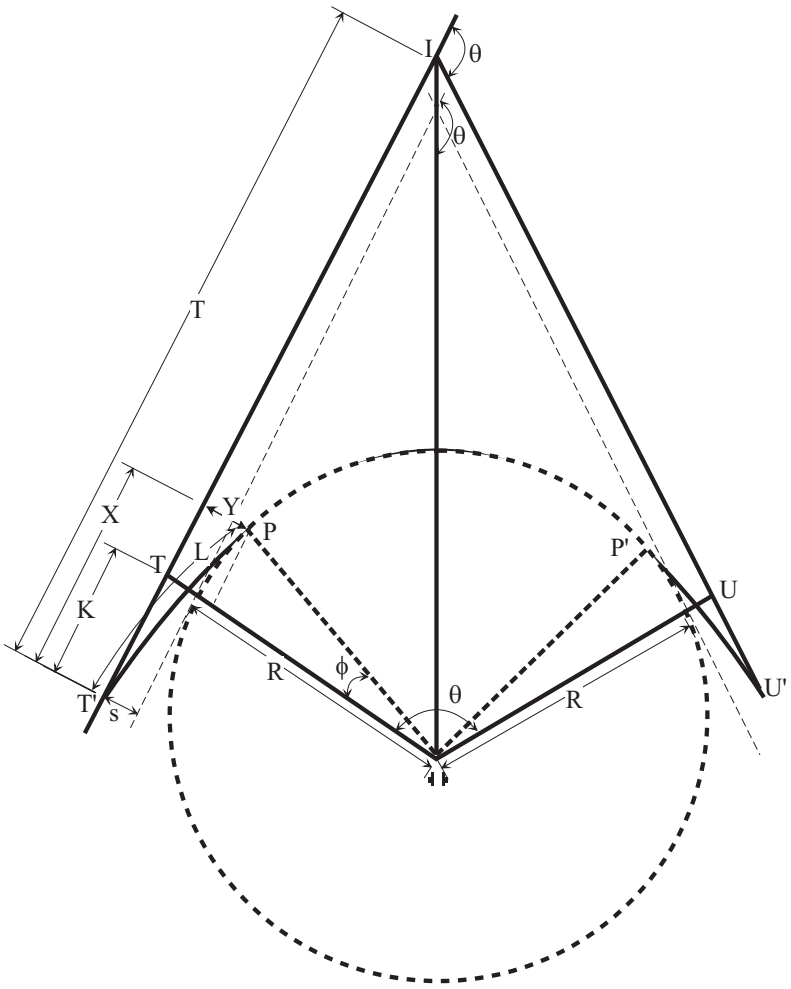

Figure 2.4*: The horizontal geometry of a transition curve*

Figure (2.3) shows the geometry of a transitional curve consisting of an entry transitional curve of decreasing radius of curvature, a circular curve of constant radius of curvature and an exit transitional curve of increasing radius of curvature. The effect of introducing a transitional curve is to move the circular curve inwards by the amount i.e. the distance between main tangents and theoretical tangents to the circle parallel to the main tangents. K is the amount by which the spiral curve tangent point extends past the circular curve tangent point also called spiral extension.

For a small increment in the spiral length dl, the offset (coordinates) differences (dx, dy) to any point on the spiral can be determined as

$$dx = dl\cos\phi \quad \text{and} \quad dy = dl\sin\phi \tag{2.16}$$

$$x = \int \cos\phi \cdot dl \quad \text{and} \quad y = \int \sin\phi \cdot dl \tag{2.17}$$

Equation (2.17) has no closed solution and are expanded using the Taylor's series of the $\cos\phi$ and $\sin\phi$ functions. The offsets x and y are then computed by integrating the expanded series term by term as shown in Equations (2.18) through (2.21) (Ressel,1999).

$$x = \int \left[1 - \frac{\phi^2}{2!} + \frac{\phi^4}{4!} - \cdots\right] \cdot dl = \int \left[1 - \frac{\phi^2}{2} + \frac{\phi^4}{24} - \cdots\right] \cdot dl \tag{2.18}$$

$$y = \int \left[\phi - \frac{\phi^3}{3!} + \frac{\phi^5}{5!} - \cdots\right] \cdot dl = \int \left[\phi - \frac{\phi^3}{6} + \frac{\phi^5}{120} - \cdots\right] \cdot dl \tag{2.19}$$

Substituting Equation (2.14) into Equations (2.18) and (2.19) then integrating, one obtains

$$x = \int \left[1 - \frac{1}{2}\left(\frac{l^2}{2RL}\right)^2 + \frac{1}{24}\left(\frac{l^2}{2RL}\right)^4 - \cdots\right] \cdot dl = \int \left[1 - \frac{l^4}{8R^2L^2} + \frac{l^8}{384R^4L^4} - \cdots\right] \cdot dl \tag{2.20}$$

$$y = \int \left[\frac{l^2}{2RL} - \frac{1}{6}\left(\frac{l^2}{2RL}\right)^3 + \frac{1}{120}\left(\frac{l^2}{2RL}\right)^5 - \cdots\right] \cdot dl = \int \left[\frac{l^2}{2RL} - \frac{l^6}{48R^3L^3} + \frac{l^{10}}{3840R^5L^5} - \cdots\right] \cdot dl \tag{2.21}$$

The offsets are therefore determined using the following equations

$$x = l - \frac{l^5}{40R^2L^2} + \frac{l^9}{3456R^4L^4} - \cdots = l - \frac{l^5}{40A^4} + \frac{l^9}{3456A^8} - \cdots \tag{2.22}$$

$$y = \frac{l^3}{6RL} - \frac{l^7}{336R^3L^3} + \frac{l^{11}}{42240R^5L^5} - \cdots = \frac{l^3}{6A^2} - \frac{l^7}{336A^6} + \frac{l^{11}}{42240A^{10}} - \cdots \tag{2.23}$$

The total spiral angle Φ subtended at full length of the spiral L is given by

$$\Phi = \frac{L}{2R} \tag{2.24}$$

while the total offsets at the full length of the spiral are given by

$$X = L - \frac{L^3}{40R^2} + \frac{L^5}{3456R^4} - \cdots \tag{2.25}$$

$$Y = \frac{L^2}{6R} - \frac{L^4}{336R^3} + \frac{L^6}{42240R^5} - \cdots \tag{2.26}$$

respectively. The spiral extension K can therefore be determined as

$$K = X - R\sin\Phi = X - \left[\frac{L}{2} - \frac{L^3}{48R^2} + \frac{L^5}{3840R^4} - \cdots\right] = \frac{L}{2} - \frac{L^3}{240R^2}\cdots \tag{2.27}$$

and tangent length T as

$$T = (R + s)\tan\left(\theta/2\right) + K \tag{2.28}$$

The offset distance from the tangent of the spiral to the tangent of the circular curve s (*cf.* Figure 2.3) is given by the following Equation (2.29).

$$s = Y - R(1 - \cos\Phi) \tag{2.29}$$

The three main distinguishing characteristics of the horizontal curve alignment to be investigated are as outlined in Table (2.3).

Table 2.3: *Characteristics of track horizontal alignment*

	Straights	Circular Curves	Clothoid (Transition Curves)
Curvature	$\kappa = 0$	$\kappa = \frac{1}{R}$	$\kappa = \frac{d\phi}{dl} = \frac{1}{r} = \frac{l}{LR}$

In Germany, the clothoid has been adopted as the standard design of all transition curves due to computation practicability and the proportionality between the curvature $1/R$ and the curve length L (Ressel, 1999). The clothoid parameter serves as the curvature factor. From the geometry of the clothoid, a change in the clothoid parameter A by a known factor results in change in lengths when all directions and proportion values are unchanged.

2.2.2 Cross section elements

The selection and configuration of the road cross section is also a major focus in road design. The cross section relates directly to the number of lanes being provided, the road width and location of shoulders, slopes, embankments and ditches. The design standards and regulations for the road cross-sections in Germany are specified in RAS-Q (*cf.* Section 2.1) and will not be discussed in this work.

2.2.2.1 Superelevation

As the radius of curvature is shortened, the centrifugal force on the vehicle increases radially outward. The raising of the outward side of the roadway to counter balance the increase in centrifugal force is referred to as superelevation. Superelevation is therefore the constant or constantly decreasing or increasing track bed roll angle provided to counter balanced the centrifugal force. The safe speed at any point on the curve is determined by radius of curvature, the side friction and the rate of superelevation. The highest practical superelevation is usually selected for the horizontal curve design and is about *6%* for rural or urban freeways. From the equations for circular motion, friction and inclined plane relationship, the superelevation can be estimated using Equation (2.30) (AASHTO, 1994).

$$e = \frac{v^2}{127R} - f \tag{2.30}$$

where e is the maximum superelevation rate in percentage *(%)*, f s the maximum allowable friction factor i.e. side slipping friction factor, v is the speed in *km/h* and R is the radius in metres.

2.2.3 Vertical alignment

Objectives of vertical curve design are to determine elevation that ensures proper drainage and offers acceptable level of safety. This is achieved by vertical curves which join two road or rail gradients and may be convex (summit) or concave (sag) so as to have constant change in gradient and to provide adequate sighting distance. Several factors influence the vertical alignment and these include topography, stopping sight distance, traffic volume and road cross section. As per the AASHTO (1994) guideline, vertical curves are always parabolic and

symmetrical. Parabolas are used in the design of vertical curves due to their simplicity. The elements of vertical design parameters include straight profile lines connected by vertical parabolic curves, mainly due to easier computation of vertical offsets and smooth transition from tangent to curve and back to tangent. The change in centreline of the road profile indicates positive or negative grade. An ideal vertical grade is a situation in which the cut is balanced against the fill without a deal of borrows or an excess of cut material to be wasted (Wright, 1996). Similar to horizontal road alignment, vertical alignment should follow the natural terrain.

Grade and grade controls have a significant effect on the operating costs of vehicles by increasing fuel consumption, emissions and noise while reducing speed. The objective is therefore to minimise the operating cost. Depending on the surrounding topography, class of road and the design speed, the grade is usually selected to maximize on cuts and fills and are limited to below *12%* for safety reasons. In tunnels for example, grades of less than *2.5%* are recommended.

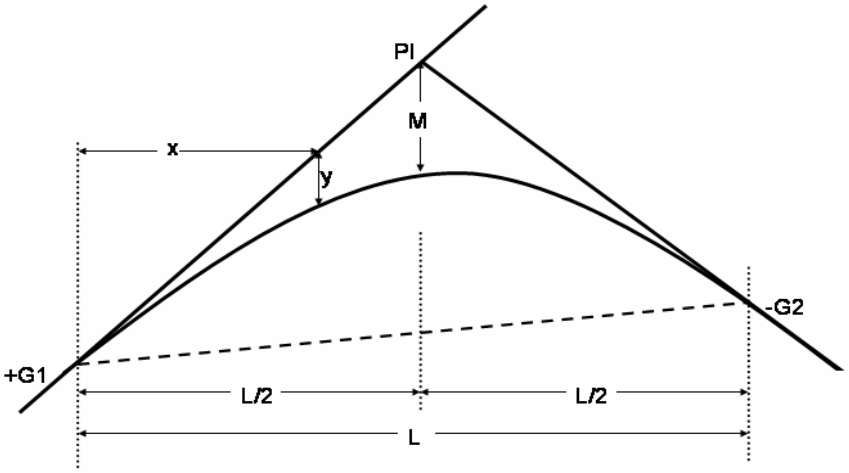

Figure 2.5: *Vertical curve elements (After Wright, 1996)*

Figure (2.5) represents vertical curve elements where $G1$ is the grade of the first tangent line and $G2$ is the grade of the second tangent line both intersecting at point PI. M is the mid-ordinate of the parabola joining the two tangents while L is he horizontal separation between the start and end of the curve. If E_x is the roadway elevation at a point x metres away from the beginning of the curve, the parabola ensures constant rate of change of slope so long as the tangents are equal as shown in Equations (2.31) through (2.33).

$$E_x = ax^2 + bx + c \tag{2.31}$$

$$\frac{d^2E_x}{dx^2} = 2a = \frac{G2 - G1}{100L} \tag{2.32}$$

$$a = \frac{G2 - G1}{200L} \tag{2.33}$$

The design controls of a vertical curve is a ratio of the rate of curvature K, as shown in Equation (2.34), per percentage of the algebraic difference between the two grades A where

G1 and G2 are the grades, E_{G1} is the elevation at the beginning of the curve, E_x is the elevation at a point on the curve x metres from the start of the curve.

$$K = \frac{L}{A} \tag{2.34}$$

and

$$A = G2 - G1 \tag{2.35}$$

The offset of the curve from the initial grade line y at any point of the curve is given by

$$y = \frac{(G1 - G2)x^2}{200L} = \frac{Ax^2}{200L} = \frac{x^2}{200K} \tag{2.36}$$

While the elevation of a point on the curve E_x located at a distance x from the start of curve is given by Equation (2.37)

$$E_x = E_{G1} + G1x + \frac{(G2 - G1)x^2}{200L} = E_{G1} + G1x + \frac{Ax^2}{200L} \tag{2.37}$$

The grades also affect the stopping distance by shortening the distance on positive grades and lengthening the distance on negative grades. These formulations are only applicable in symmetrical parabolic curves in which the tangents are assumed to be of equal length. Although horizontal and vertical alignments are mutually related, the vertical curves and cross section have only been included in this work for completeness and have no relevance to this work.

2.2.4 Stopping sight distance

Sight distance is the length of the roadway visible to the driver. Sight obstructions such as wall, cut slopes and buildings limit the sight distance on the curves. Adjustment to the alignment is therefore required to provide adequate sight distance since most of the obstructions are rarely removable. Stopping sight distance is the sum of brake reaction distance i.e. the distance traversed by the vehicle after the driver sights an object to the instant the brake is applied and the breaking distance i.e. the distance needed to stop the vehicle after the brake is applied. The stopping sight distance is dependent on the height of the driver's eye above the road surface, the specified object height above the road surface and the lateral position of the sight obstructions within the driver's line of sight. Figure (2.6) provides a formula for evaluation of the total sight distance on horizontal curves based on a maximum permissible superelevation of *6%*, height of eye of *1.07 m, 0.61 m* height of object for stopping and reaction distance. The stopping sight distance is given by AASHTO (1994, pp 117-118) as shown in Equation (2.38)

$$SSD = BRD + BD$$

$$SSD = 0.278vT + 0.039\frac{v^2}{a} \tag{2.38}$$

where

SSD - Stopping Sight Distance *(m)*
BRD - Break Reaction Distance *(m)*
BD - Break Distance *(m)*
v - is the design speed of the vehicle in *Km/h*
a - is the deceleration rate ($3.4m/s^2$)

T - is the AASHTO assumed break reaction time of *2.5 seconds*.

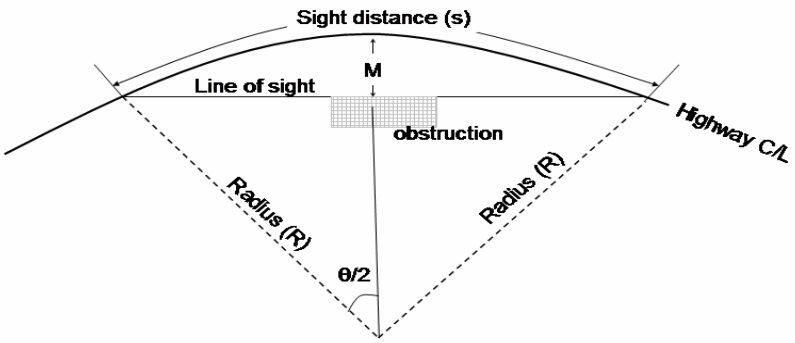

Figure 2.6: *Sight distance measurement for horizontal curves*

With R is the radius in metres and θ as the deflection angle in degrees. The mid-ordinate M and stopping sight distance S are computed as shown in Equations (2.39) and (2.40) respectively.

$$M = R\left(1-\cos\left(\frac{\theta}{2}\right)\right) \tag{2.39}$$

$$S = \frac{\theta\pi R}{180} \tag{2.40}$$

Stopping sight distance is a determinant factor in horizontal curve design, in that, the minimum length of a curve or tangent must be equal or longer than the stopping sight distance hence determines the minimum radius of curvature. This permits the driver to see an obstacle soon enough to stop. Similarly the minimum length of the crest vertical curve is also based on sight distance.

3 Sensors and Systems for Land-Based Positioning and Navigation

Equation Section 3

Several sensors have been used as independent navigation instruments for providing either the position, the course or the distance travelled by a vehicle in military, roads, railways and other industrial engineering applications. The advancements made in electronic industry and the availability of digital maps covering larger regions has widened the demand for more accurate and robust land-based positioning and navigation systems resulting in the integration of sophisticated, real-time, multi-tasking, multi-sensor systems. These integrated units and systems supplement each other in curbing and reducing the inherent sensor errors. Emphasis is currently being laid on the suitability of the low-cost sensors and systems in achieving the navigation performance. This section briefly discusses the different sensors and systems used in land-based applications, their concepts, requirements and their susceptibility to different error sources.

3.1 Inertial sensors and inertial navigation systems (INS)

Inertial sensors and inertial navigation systems measure rotational and linear movements without external reference and have been used in navigation systems to provide vehicle guidance information and lately in industrial robots (Gao *et al.*, 2003).

3.1.1 Types of inertial sensors and INS

The inertial sensors widely used in land-based applications today are the accelerometers and gyroscopes. Presented in this section are examples of the gyroscopes such as spinning wheel gyros, fibre-optic gyros, laser gyros (including MEMS gyros) and examples of the accelerometers such as guidance and vibratory accelerometers (including MEMS accelerometers). Inertial measuring units and inertial navigation systems including their error sources are also discussed. The developments in solid-state accelerometers and gyroscopes have led to the availability of small-sized low-cost inertial sensors and inertial navigation systems.

3.1.1.1 Gyroscopes

Gyroscopes or gyros as they are commonly known measure the angles or angular rates normally given in degree per second *(°/s)* or degree per hour *(°/h)* about a single axis. The gyros can be categorised into several groups based on the operating principles, examples of which include spinning mass gyros, optical gyros and vibrating gyros.

Mechanical gyros function on the basis of a rapidly *spinning mass* (Jekeli, 2001) and consist of mass spinning steadily about a free movable axis. They are based on the principles of conservation of angular momentum. A spinning mass gyro will therefore exhibit the following characteristics:

- the gyro will maintain a constant angular momentum in inertial space as long as the gyro is kept in a torque free environment. This is the principle used by systems which are space-stabilised by gyroscopes.

- the gyro will change direction of the sensitive axis in a well defined way when a torque is applied to it in a controlled manner.
- the torque proportional to the resulting change in gyros angular momentum will be introduced when the gyro assembly is rotated. This is the principle used by the strap-down systems.

Although spinning mass gyros have been proved by research to be capable of achieving high accuracy through sophisticated mechanical design with low friction bearings, the high cost of the hardware needed to support the spinning mass has limited their use in navigation.

Optical gyros operate on the principle that for a beam of light reflected around steadily in an enclosure, the duration between the moments of light emission to eventual reception will be different, when the enclosure is rotated. There are two types of optical gyros, the ring laser gyro and the fibre-optic gyros. In ring laser gyros (RLG), the laser beam reflection is through normal glass mirrors while in fibre-optic gyros (FOG) the transmission and reflection is through fibre-optics. The rotational sensitivity of the optical gyros is based on the *Sagnac effect* principle. The ring laser gyro measures the change in path length ΔL as a function of the angular velocity of rotation Ω, radius of the circular beam path r and the speed of light c as shown in Equation (3.1) (Kelly, 1996)

$$\Delta L = \frac{4\pi r^2 \Omega}{c} \tag{3.1}$$

The FOGs on the other hand measures the phase difference for a single round trip which grows linearly with the number of turns in a fibre-optic light pipe as given in Equation (3.2) (*ibid*.).

$$\Delta\phi = \frac{4n\pi Lr\Omega}{c\lambda} \tag{3.2}$$

where $\Delta\phi$ is the phase difference, n is a constant and λ is the wavelength. The FOGs are solid-state devices i.e. no moving parts and therefore have several advantages over RLGs in that they are simpler, more reliable, of relatively lower cost and require no critical mirrors. A dominant source of noise is the interference between the signal wave and the backscattering in the resonator. This phenomenon which causes output fluctuations or output nonlinearity at low rotation rate is referred to in the latter case as lock-in phenomenon. RLGs are known to be independent of this source of noise. Optical gyros, in general, are not influenced by linear acceleration as compared to the mechanical gyros and have successfully been used in strapdown navigation applications. A most recently emerging technology in optical gyros are the interferometric FOGs (IFOGs) which provide the closed optical path by a multi-turn optical fibre coil wound on another coil. They are more compact and have relatively lower cost than the RLGs (Smith and Weyrauch, 1990).

3.1.1.2 Accelerometers

An accelerometer is used to measure the specific forces (not acceleration) acting on a device along a sensitive axis. This specific force consists of kinematic acceleration of the device with respect to inertial space (conservation of linear momentum) and gravitational acceleration acting on the device which is caused by the mass distribution within the earth and masses of other celestial bodies. Accelerometers that are used to sense the reaction to vehicular acceleration with respect to inertial frame are also referred to as specific force sensors (SFS).

Acceleration denoted by the vector a, is the total Newtonian force F acting on an object divided by its mass M as given in Equation (3.3) (Hou, 2004)

$$a = \frac{F}{M} \tag{3.3}$$

Just like the gyros, several types of accelerometers exist and vary from the traditional spring-mass devices to modern vibrating beam quartz or silicon technologies. Accelerometers can also be classified into two categories, depending on their application,

- *Guidance accelerometers* are used to measure steady state accelerations of rigid bodies, for example acceleration of an automobile.
- *Vibratory or seismic accelerometers* are used to measure sinusoidal accelerations for example vibration in structures, earthquake and tsunami detection.

3.1.1.3 MEMS inertial sensors

A relatively modern technology that is currently revolutionizing the inertial sensor construction and performance is the *micro-electro-mechanical systems* (MEMS), usually designed as an electronically driven resonator and fabricated in a single piece of quartz (*quartz rate sensors*) or silicon (*silicon tuning fork sensors*). Advantages of MEMS inertial sensors over other inertial sensors include: simple construction since only one piece is required, small size and weight, modest power consumption, lower cost at mass production and relative durability. The physical mechanism underlying the MEMS sensors are piezo resistive, electromagnetic, piezo electric, optical, tunnelling and capacitive (Bernstein, 2003).

MEMS gyros: all MEMS gyroscopes use the Coriolis effect on the vibrating resonator chip to sense the rate of rotation. The different types of MEMS gyros fall into four categories. *Tuning-fork* gyros consist of a pair of masses that are driven to oscillate at equal amplitude but in opposite directions. *Vibrating (oscillating) wheel* gyros consist of a wheel driven to vibrate about its axis of symmetry and a rotation about either in-plane axis results in the wheels tilting. Other types are *wine glass resonator gyros* and *Foucault pendulum gyros*. Advantages of rate sensors in general include resistance to external shock, wide range of operating temperature *(-40°C* to *85°C*), low drift and high reliability. Drifts which are as low as *1°/h* (temperature dependent) have been proved (Barbour and Schmidt, 2001). MEMS gyros have attracted and have been applied, in companion with MEMS accelerometers, in providing attitude and heading information for navigation purposes, automotive ride stabilisation, camera stabilisation, virtual reality, inertial mouse for computers and robotics.

MEMS accelerometers:
Similar to MEMS gyros, the MEMS accelerometer are also produced using the principle of integrated circuits and have the advantages of being small, relatively cheap as well as being robust and reliable. The only setback with MEMS sensors, both gyros and accelerometers, is that they are yet to achieve the accuracies achievable by other inertial sensors. *100 μg* bias error and *100 ppm* of scale factor error has been achieved but further improvements are expected (Barbour and Schmidt, 2001).

Optical MEMS sensors are also called micro-optical electromechanical systems (MOEMS) sensors. Optical switching technology used in data communication is now being researched for potential combination of micro-optics with micromechanics to create a broader use of devices called MOEMS.

3.1.1.4 Inertial measuring unit (IMU)

An inertial measuring unit, the so called IMU, consists of a triad accelerometers, a triad gyroscopes and support electronics which basically comprise of a digitization circuitry and a CPU for signal conditioning and temperature compensation (*cf.* Figure 3.1). The three accelerometers and three gyroscopes mounted such that their measuring axes are orthogonal to each other offering six-degrees of freedom. The system can either be gimballed, suspended in fluid or strapdown system. IMU provides useful information about the platform on which it is mounted i.e. the acceleration, rotation and orientation. Although an IMU provides these information about the platform on which it's mounted, calculations are required to transform the sensed accelerations and angular rates into navigationally useful information such as position, velocity and attitude. The basic components of the inertial measuring units i.e. gyros and accelerometers, have remained essentially the same over the years but have improved in performance and acquired a wide range of applicability.

3.1.1.5 Inertial navigation systems (INS)

Navigation state is the position, velocity and attitude of an object expressed as time variable function. Therefore a combination of sensors cable of deriving the navigation state is a navigation system (El-Sheimy, 2003; Hou, 2004). Sensor fusion or integration is the combination of information from different sensors to obtain the best estimate of the parameters not measured directly through detection of the different sensors failures and/or noise. In land based applications, this is made possible by the capability of GPS positioning e.g. car navigation systems, lane keeping systems and autonomous ground vehicle in military.

An inertial navigation system (INS) comprises of an IMU, other aiding sensors, a platform on which they are mounted, a stabilization mechanism and a software to perform navigation computation (Jekeli, 2001). This involves solution of a large set of differential equations to convert these measurements into position and attitude estimates starting with a known initial position and orientation. Figure (3.1) shows the different steps performed by the INS to estimate the position, velocity and attitude of the vehicle. The inertial sensor measurements from the three gyroscopes and accelerometers are time integrated to obtain body's position and velocity. The IMU results are then fed into a Kalman filter and corrected for inertial sensor errors to obtain best estimated solution. An INS is therefore an integrated system consisting of a detector and an integrator that are used to detect acceleration and then integrate it to derive velocity, integrate the velocity to derive displacement (Schwarz and El-Sheimy, 1999).

An appropriately initialised inertial navigation system is capable of continuous determination of vehicle position and velocity using the properties of inertial sensors mounted aboard the vehicle. For this to be achieved, four functions must be performed by the inertial sensors;

- Instruments a reference frame usually accomplished by gyroscopes and the supporting electronics i.e. determine a coordinate frame for the accelerometers
- specific force measurement done by use of accelerometers
- gravitational force modelling accomplished using a gravity field model
- integration of the specific force data over time to obtain velocity and position accomplished by onboard digital computers (Lightsey, 2006).

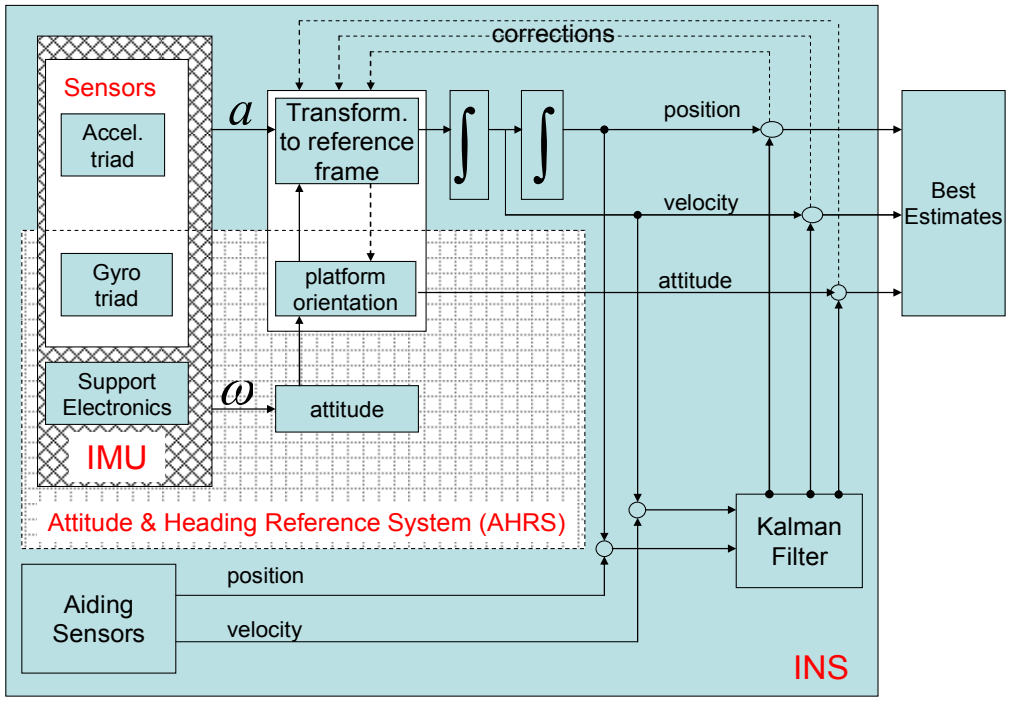

Figure 3.1: *Inertial sensors, IMU and INS* (After El-Sheimy, 2003, modified).

Inertial measuring unit are usually incorporated into navigation systems due to their independence and subsequent immunity to jamming. In general, INS solves the second order differential equation resulting from Newton's second law of motion as represented by Equation (3.4) below (Lightsey, 2006)

$$\frac{\partial^2 r^I}{\partial t^2} = C_P^I f^P + G^I \tag{3.4}$$

where

$\frac{\partial^2 r^I}{\partial t^2}$ is the inertial referenced acceleration vector of the vehicle

C_P^I is the direction cosine matrix (DCM) derived from a gyroscope and used in transforming the acceleration vector from navigation platform to the inertial reference frame

f^P is the specific force relative to the measured acceleration

G^I is the gravity acceleration at the current location

while subscripts I and P represents inertial and platform reference frames respectively. The velocity and position are obtained by performing single and double integration respectively of the measured acceleration with respect to time.

3.1.2 Inertial sensors, IMU and INS errors

All onboard sensors and systems are affected by deterministic and random measurement errors. Different sensors are prone to environmental factors, manufacturing and system set-up errors. Due to these, inertial sensors for example, experience bias error, scale factor errors, gain error, noise and other random effects. The errors build up over time corrupting the measurement data. For accurate information to be derived from the measurements, the contamination noise and displacement errors must be removed (Gao *et al.*, 2003). Inertial

sensors and IMUs performance is characterized by the stability of the inherent biases in the output and scale factor errors (Jekeli, 2001).

High accuracy INS equipped with expensive IMUs are usually calibrated by the manufacturers in the laboratory and thus no extra calibration is required. However low-cost INS and dead reckoning (DR) systems on the other hand require calibration every time they are to be put into use. This is usually done by averaging raw angular velocities obtained from the gyroscope over a period of *15-20* minutes in static mode (Nayak, 2000).

3.1.2.1 Gyroscope errors

Two main error sources that affect the accuracy of the gyro measurements are the bias and gyro scale factor errors. The gyro bias usually measured in *(°/s)* or *(°/h)* is the increase in angular error over time and contributes an offset error in the output. The gyro bias can estimated every time the sensor is static then subtracted from the subsequent measurements. The gyro bias is a temperature sensitive variable error. Therefore, due to temperature variation and other environmental factors, the unaided gyro will read a non-zero value even if the angular velocity is zero. If the gyro is aided with GPS or odometer, the heading drift can be corrected. The errors due to gyro bias which increases as the square of the distance travelled are modelled using a random walk process (modelled as fixed and random). Chung *et al.* (2001) also reports that trivial short calibration method eliminates bias drift i.e. calibration prior to each measurement mission.

Gyro scale factor error is the non-linearity of the scale factor and is an indication of angular error which occurs during rotation given in parts per million or in percentage. The error affects the gyro measurements only when the vehicle is taking a turn. The gyro scale factor error is calibrated by rotating the gyro in a varying but precisely controlled input rates and then obtaining reading differences. The non linear perturbations in the scale factor are though typically very small relative to gyro bias (drift) errors and are therefore insignificant (Gat, 1995).

Gyro noise when integrated can be modelled as Random-Walk. The gyro noise influence on position and curvature estimation is investigated further in Section (5.3.3). Another important error source is due to the temperature. Most sensors model the temperature effects and compensate for temperature influences internally but this is never perfect. Temperature dependence of the bias and scale factor errors complicates the calibration of the sensors.

Misalignment refers to mechanical misalignment between the gyro or accelerometer measurement axes. It is defined as a random constant and sometimes referred to as non-orthogonality error. There are several other errors which are specific to the different types of gyroscopes and are listed in details in Titterton and Weston (1997).

3.1.2.2 Accelerometer errors

The accuracy of an accelerometer depends on the type, the resolution, the range of measurable acceleration, bias, scale factor error and the sensitivity to temperature change. Analogous to gyro, accelerometers performance are also determined using bias (m/s^2), scale factor error and noise. The error in position due to accelerometer bias is proportional to the square of the time (t). Another source of error in accelerometers is the misalignment where the measuring axis is not parallel to the intended axis. The specific force $f = g - a$ is reduced when the

measuring axis is at angle to the vertical axis hence recording reduced reading (Forder, 2000). Gao (2002) gives the general error equation for the accelerometer measurements $\tilde{a}_x$, along the x-axis direction as shown in Equation (3.5)

$$\tilde{a}_x = a_x + b_a + s_x a_x + G_c\theta + c_y a_y + c_z a_z + \varepsilon \tag{3.5}$$

where a_x is the true acceleration, b_a is the bias error, s_x is the scale factor error, G_c is the gravitational constant *(9.81m/s²)*, θ is the pitch angle, c_y, c_z are the cross coupling factors that depend on mounting and ε is the random noise and vibration in the sensor signal. The bias and scale factor errors can be reduced substantially by performing individual calibration on the accelerometer and gyroscopes (Chung *et al.*, 2001). Gao *et al.* (2003) state that two methods can be used to eliminate accelerometer error; bias error modelling and wave form correction. The remaining systematic and random errors can be eliminated using a Kalman filter process (*cf.* Figure 3.2). Kalman filter has been used to fuse the inertial sensor measurements to other sensor data to curb the inherent gyro and accelerometer drifts and offsets. The effects of inertial sensors' noises are reduced using stochastic modelling.

3.1.2.3 MEMS inertial sensors errors

The errors of the MEMS based inertial sensors, that is MEMS accelerometers and MEMS gyro, can also be categorized into stochastic and deterministic. Two main errors affecting MEMS based sensors are bias and scale factor errors. There are also several contributors to noise in MEMS sensors such as the read out electronics, mechanical damping and electrical resistance. MEMS IMUs, are also subject to noise, bias instability, scale factor error, non-orthogonality of the sensitive axes and temperature sensitivities (bias and scale factor). MEMS based IMU currently available are laboratory tested by the manufacturers who then specify their performance but independent field and laboratory tests by the users are possible (Jöns, 2001; Niu *et al.* 2006). The errors of MEMS INS are due to drift and noise in the signals which increases rapidly with the integration processes to obtain the vehicle position and velocity. The common most method of reducing the MEMS INS errors is by use of aiding sensors such GPS.

3.1.2.4 Inertial measuring unit error sources

The sensors exhibit different error characteristics. Their performances are therefore gauged using offsets (drift), scale factor error, temperature and pressure effects, zero-rate output (ZRO) and resolution. IMU errors are mainly due to accelerometers and gyroscope errors. The IMUs are affected by cumulative errors which subsequently lead to angular and position drifts. Apart from the noise, biases, zero velocity update (ZUPT) and Schuler tuning, other issues are the synchronisation of the multiple sensors, calibration, alignment and transformation of coordinates. These errors must be reduced through frequent calibration of their systematic effects using real time software estimations procedures for example Kalman filtering. Another major error source, especially in low-cost sensors, is the temperature bias. This is usually compensated by using temperature calibration tables. There is therefore need for built-in-test capabilities to detect and eliminated some of the errors.

Coning error for the gyro and culling errors for the accelerometers due to the assumption that a strapdown platform is non-rotating while performing numerical integration. IMUs are factory calibrated by the manufacturer and error specifications provided. In summary, the IMU errors are due to the individual sensor errors such as inertial sensor bias error, scale

factor error, cross axis coupling and misalignment errors due to misalignment angle along the sensitive axes which results in sensing of the gravity component as part of acceleration (gravitational effects).

3.1.2.5 INS error sources

Inertial navigation systems are affected by three categories of errors

- inertial measuring unit output errors
- system set-up and aiding sensors output errors
- computational process errors for example due to numerical integration.

Therefore the error dynamics in an INS are mainly due to aiding sensors errors and IMU error characteristics which arise from the inertial sensors. The Schuler oscillation and the linear position error accumulation can be modelled by a system of linear differential equations whose input would be the accelerometer and gyro errors. The IMU errors can in turn be modelled using Markov processes, random walk processes or random constants. Aiding sensors errors, on the other hand, are a combination of correlated random error sources such as GPS signal ionospheric delay and multipath and uncorrelated errors such as GPS receiver internal noise. The correlated error sources are modelled by Markov process or by white noise models. The near constant errors are modelled as random constant while the uncorrelated noise are modelled as white noise.

In summary, inertial sensor and inertial navigation system errors are divided into two spectral signatures. *Short term errors* are errors that change rapidly while *long term errors* are errors that change slowly (Skaloud *et al.*, 1999). The short term errors, are inertial errors in the high frequency range and include errors due to vibrations and quantization. These errors are not directly observable and can only be suppressed prior to mechanization or can be pre-filtered if their distribution is known. The long term errors, on the other hand, are inertial errors in low frequency range and include gyro drift, accelerometer bias, scale factor error, non-orthogonality and misalignment errors. The long term errors are modelled from long term observations as error states or can be estimated from the manufacturer specifications. They are however not constant. For example the bias and scale factor errors are variables of temperature while noise is variable of the dynamic environment in which the system is being operated. The errors are also classified into two regions, *motion* and *non-motion*. Figure (3.2) shows the amplitude of the different sensor errors against the sampling frequency and their corresponding estimation and reduction concepts.

Gao *et al.* (2003) also proposes a technique called waveform integration for the estimation and reduction of the random errors. The velocity waveform is fitted by a second order polynomial and used to estimate the drift in the acceleration measurements based on the assumption that the initial and final velocities are zero. In general Kalman filtering is used to fuse the inertial sensor measurements with external aiding sensor measurements to curb and correct the errors. The purpose of the Kalman filter in the inertial navigation system is therefore to estimate and eliminate the remaining systematic and random errors.

The high frequency noise present can be filtered using a low-pass filter. The double integration of the inertial sensor measurements, for example to derive displacement from acceleration, also suppresses the high frequency noise. The long term errors are reduced by updating the filter model with the observed error state vector coming from for example GPS measurements (position and velocity).

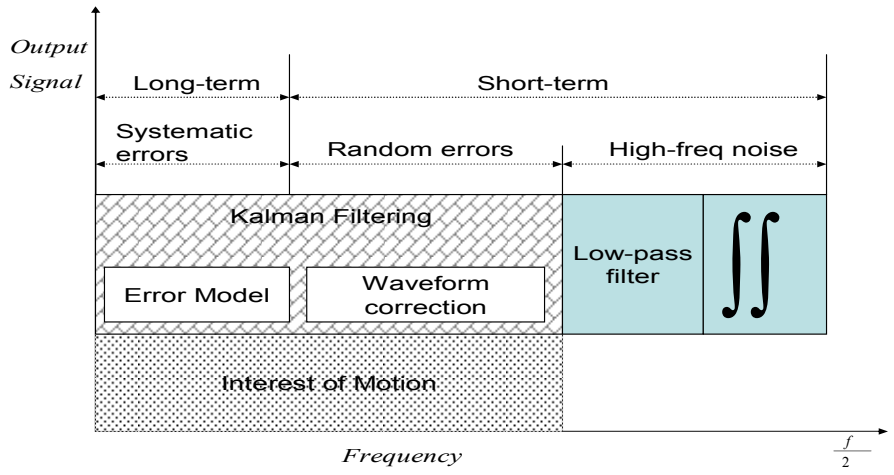

Figure 3.2: *INS error components and corresponding reduction strategies (Gao et al., 2003)*

3.2 Wheel speed sensors

The distance increments of the vehicle can be obtained in two different ways;

- from the speedometer assembly of the vehicle
- a hub of sensors attached to the wheel of the vehicle.

Wheel speed sensors are devices used for reading the speed of the vehicle's wheel rotation and are a fundamental component of the ABS. There are several techniques used in the generation and pickup of the signals indicating the speed of rotation of the wheels. These are summarised in the section below.

3.2.1 Types of wheel speed sensors

The wheel speed sensors can be classified as either *exposed* or *concealed* or *as four wheel* or *two wheel sensors*. Another classification based on the type of sensor used to detect the speed of the wheel rotation are *optical sensors, magnetoresistive sensor, inductive sensors, Hall effect sensors and variable reluctance devices.*

3.2.1.1 Optical wheel speed sensors

An optical wheel speed sensor uses a photo cell, light emitting diode (LED) and a two-blade mirror reflector to generate an electrical signal. When the vehicle is in motion, the two-bladed reflector rotates through the LED light beam, breaking the beam twice for each revolution of the reflector. Each time the reflector passes through the LED beam, light is reflected to the photo cell which then generates an electric signal.

3.2.1.2 Magnetic pickup-type

These consist of a magnet attached to the wheels and a pickup sensor attached to the frame. Each time the wheel rotates, the magnet causes the pickup sensor to generate a voltage that is then counted and multiplied by the wheel circumference to obtain distance increments. The magnet-pickup setting can be replaced by a toothed wheel and a magnetic sensor. The sensor

counts the pulses as each tooth of the wheel passes near it. The magneto resistive sensors have the advantage in that they have contact-less operation enabling a relatively longer operation life, they are highly sensitive and can operate at high temperatures. These types of sensors are further categorised as Hall effect sensors, magnetoresistive sensors, inductive sensor and magneto reluctance sensors.

3.2.1.2.1 Hall effect sensors

In the Hall effect sensors, a charge carrier (*V*) moving through a material experiences a deflection because of an applied magnetic field (F_H) resulting in a measurable potential difference across the side of the material traverse to the magnetic field and the direction of current (*cf.* Figure 3.3). The pulse or signal generated by the Hall effect sensors are proportional to the strength of the magnetic field and allows for measurement of the phase of rotation. Hall effect sensors are therefore more accurate since they are sinusoidal and senses even at lower wheel speeds.

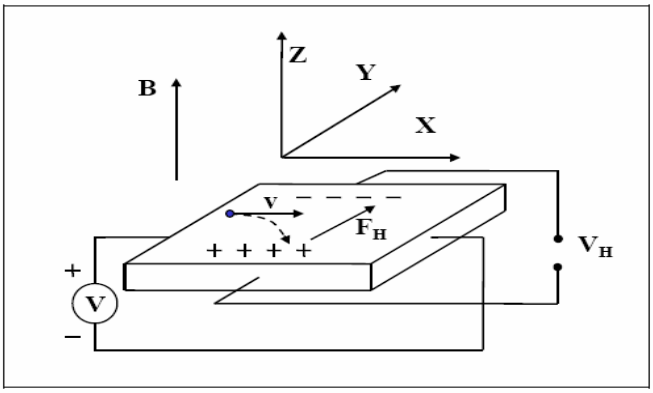

Figure 3.3: *Hall Effect* (source: Van Zeghbroeck, 1998)

3.2.1.2.2 Magnetoresistive sensors

The resistance of the sensing element is a function of the direction and magnitude (proximity) of an applied magnetic field. Magnetoresistive sensors consist of four magnetically sensitive resistors in a Wheatstone bridge configuration, with each resistor arranged to maximize sensitivity and minimize temperature influences. In the presence of magnetic field, the values of the resistors change causing a bridge imbalance and generating an output voltage proportional to the magnetic field strength.

3.2.1.2.3 Inductive sensor

The sensors function on the principle of magnetic induction and consist of a magnet mounted in close proximity to a notched reference ring that spins at wheel speed causing a low voltage rhythmic pulse to be generated by the magnet. The computer reads the signal as wheel speed (Coconis, 2005). In an inductive sensor an oscillator circuit generates an electromotive force (e.m.f) that radiates from a ferrite core and coil assembly. The field is directed at the sensor face. When a metal target enters the field, eddy currents are induced into the surfaces of the target. This causes a reduction in the amplitude of the oscillator circuit (change in inductance).

3.2.1.2.4 Variable reluctance sensors

Variable reluctance sensors are also known as linear or rotary variable differential transformers (LUDTs or RVDTs). Variable reluctance sensors are based on Faraday's law of induction which states that *"the induced voltage or electromotive force is equal to the rate at which the magnetic flux through the circuit changes"* A simple diagram of the variable reluctance sensor is shown in Figure (3.4). As the drive-shaft rotates, the protruding tab (gear tooth) passes through the magnet causing a drop in magnetic flux and the coil perpendicular to varying magnetic field will generate voltage /current. The rate of change of the magnetic flux is proportional to the speed of rotation of the shaft. If the rotation speed of the shaft is to slow, weak or no signal is produced. This minimum rotation speed at which the signal is no longer generated is called the *threshold speed*. Note that the number of the protruding tabs mounted on the wheel accounts for the number of signals or pulses generated per wheel turn. Variable reluctance sensors are the widely used in many application because they are relatively small is size, are typically self-generating i.e. requires no external power and have low maintenance cost. The disadvantage is that they work only at medium to high speed but for certain system such as ABS zero speed sensing is a requirement i.e. below *7 km/h*.

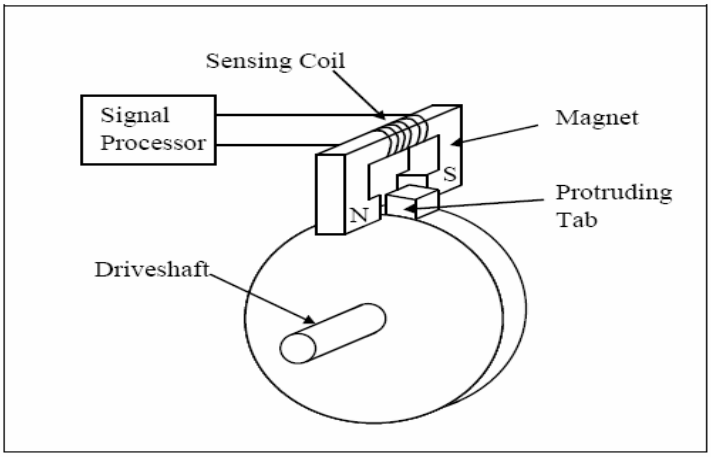

Figure 3.4: *Variable reluctance rotation sensor* (Source: Ribbens, 1994)

3.2.1.3 Magneto-optical sensors

The magneto-optical wheel speed sensors emit light signals for example infra red, which are reflected back by a nearby surface to a collector. The collector which is a silicon photodiode will induce an output signal/current proportional to the flux of the reflected light. In the absence of a magnetic field, the sensor head transmits back all the incident light intensity and reflects virtually no light intensity when full magnetisation is reached.

Anti-lock brake system (ABS)
The anti-lock brake system, the so called ABS from the German name "*Antiblockiersystem*", was designed to maintain control and stability of the vehicle during emergency breaking. The ABS monitors four wheel speed sensors to evaluate wheel slippage. The wheel speed sensors are a major component of ABS system, the other components being the *electronic control unit* (ECU), the *hydraulic control unit* (HCU or pump), the *relay* (wiring) and the *wheel speed sensors*. ABS systems can be divided into three basic categories: *four-wheel/four-channel, four-wheel/three-channel, and two-wheel (rear only)*. *Four-wheel/four-channel* uses an individual wheel speed sensor at each of the four wheels. This allows the ECU to control

braking at any one of the four wheels individually. *Four-wheel/three-channel* uses individual wheel speed sensors on each front wheel and one sensor for the rear wheels of the vehicle, which is mounted on the differential housing, to monitor rear axle speed. This system allows the ECU to control braking at either one of the front wheels individually, but at the rear wheels the ECU controls braking at both rear wheels simultaneously. The last type of system, *two-wheel*, or *rear ABS*, has only one sensor mounted on the differential housing. In this system, the ECU controls the rear brakes only, and just like in a four-wheel/three-channel system, does it simultaneously.

Variable reluctance wheel speed sensors are the most popularly used in ABS systems because of their low cost and reliability. As mentioned earlier, these type of sensors are not reliable at low speeds of about *0.45* to *1.34 m/s* (Zhao, 1997). The ABS has undergone transformation from the time Bosch filed a Patent for the invention in *1936*, the most relevant being the reduction in weight from about *6.3 kg* in *1936* to *1.6 kg* in *2003*. ABS are installed in about two thirds of the vehicles world wide (Christiani, 2005) which justifies their adoption as low-cost odometers or tachometers.

3.2.2 ABS wheel speed sensor errors

Sensor Sensitivity: The sensor sensitivity depends on the sensor type being used. The wheel speed sensors can be classified in two broader categories i.e. either as passive or active sensors. The active sensors use the principle of Hall-effect by converting wheel motion into electrical signals and perform extremely well in low-speed (Hay, 2005). On other hand, passive sensors uses the principle of variable reluctance in which sensor teeth travel through a passive magnetic field at sufficient speed to generate a low-voltage analogue waveform. Figure (3.5) shows a passive wheel speed sensor indicating the signal wave analogue converter and the sensor teeth which determines the sensor resolution.

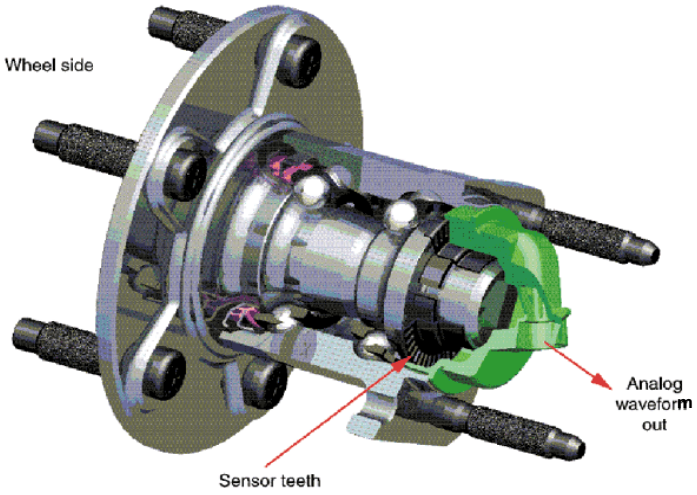

Figure 3.5: *Passive wheel speed sensor* (after Hay 2005)

Below the *threshold speed*, the wheel sensors do not generate signals strong enough to be picked up the ABS computer or the computer is programmed to ignore the signal. The dead reckoning software must therefore be designed to identify and eliminate threshold speed.

Sensor teeth count: another source of error is the sensor teeth count which is a constant in the estimation of vehicle velocity and yaw rates (*cf.* Section 3.2.2) .The higher the sensor teeth the higher the resolution.

Wheel radius estimation and tire size variation: This is mainly caused by tire pressure variation, temperature variation, altitude variation and treads wear. The tire diameter will vary as a consequence and so is the distance travelled per wheel sensor pulse.
Sampling rate: minimum rate of sampling determine the resolution of the wheel speed sensor. The ABS data and other aiding measurements also need to be synchronised.
Noise: the sensors measurements are contain noise which varies with the environment.

3.2.2.1 Calibration of the ABS wheel speed sensors

The odometer converts wheel revolution into distance through initial calibration in a process known as odometer scale factor determination. This involves for example driving the vehicle over a known distance. But due to skidding, wheel slipping, tire pressure variation, tire wear and vehicle speed, the scale factor is subject to vary. If uncompensated for, the scale factor error accumulates and subsequently results in significant position error. The odometer scale factor error is modelled using first order Gauss-Markov process.

Calibration of ABS wheel speed sensors consist the determination of the parameters required to reconstruct the vehicles' relative position i.e. distance and heading differences from the measured pulse differences. The absolute vehicle position in then derived starting at a known configuration through discrete time integration of the wheel displacement in a process called differential odometry. The calibration can be done by least-squares technique where the parameters used in the calculation of the vehicle position from the wheel measurements are determined in real time or through post processing. This can be done *geometrically* where a pre-defined trajectory is executed and processed, or *numerically* by incorporating multi-sensors such as onboard gyro and a Kalman filter algorithm. The distance increment $\Delta\tilde{d}$ travelled by the vehicle wheel can be estimated from the measured signal differences Δp as shown in Equation (3.6)

$$\Delta\tilde{d} = s\Delta p + \frac{e}{100} s\Delta p \tag{3.6}$$

where s is the scale factor and e is the scale factor error specified in percentage of the full scale.

3.2.2.2 Wheel speed sensors resolution

Wheel speed sensor resolution is basically the teeth count or signal pulses count per wheel rotation and is a function of the wheel radius or diameter. That is to say, the wheel speed sensors resolution is the minimum distance that can be detected during the distance measurement. Assuming a pulse signal count per turn n and a wheel radius R, then the smallest detectable distance $\Delta\delta$ by the wheel speed sensor is analogous to *1-pulse-signal* and is given by Equation (3.7) below

$$\Delta\delta = \frac{\text{wheel diameter}}{\text{no. of pulses per turn}} = \frac{2\pi R}{n} \tag{3.7}$$

For a wheel of radius *33 cm* and wheel speed sensors of *52 teeth*, the wheel speed sensors resolution is about *4 cm*. From the equation, one can see that the larger the number of pulses measured per wheel turn i.e. the number of sensor teeth, the better the resolution assuming the wheel radius is constant.

3.3 Other external aiding sensors

Other motion sensors that are used in navigation include Doppler velocity sensors (DVS), laser ranging devices, barometric altitude-meters, inclinometer and compasses. Barometers are used to sense change in altitude while odometers sense the distance travelled by the vehicle. Compasses are used to sense the direction of the magnetic field.

3.3.1 Compass

A compass is any device which indicates the direction of the Earth's magnetic field. A compass therefore uses Earth's magnetic field as the orientation reference by sensing the local magnetic field without aligning itself to the magnetic field of the earth. A mechanical compass allows a magnet to rotate freely in the horizontal plane and align itself to the Earth's magnetic field. A digital compass on the other hand uses a toroidal flux gate to sense the horizontal direction of the earth's magnetic field. Mechanical compasses have poor transient response and are sensitive to vehicle vibration. The mechanical compass has the advantage that it requires little maintenance and no power source. *The fluxgate compass*: is an electromagnetic saturated core device with no moving parts and consist of sensor, display and associated electronic components. The transient response is excellent and insensitive to vibrational disturbances. The Fluxgate compasses are the types usually used in mobile navigation systems. If not used with inclinometer, two axis of measurement must be incorporated. Compass errors are the variation which is the difference between true and magnetic north at any point on earth and the deviation which is caused by the interaction of the compass with the metallic and electrical systems around it. Deviation can be compensated for but never eliminated. A compass table is normally used in the elimination process.

3.3.2 Inclinometer

An inclinometer is used to measure tilt and is used in platform levelling. Inclinometer, also known as clinometer or tilt meter or gradiometer, is an instrument for measuring angles of elevation or inclination with respect to gravity i.e. horizontal angle and slope or vertical deviations. It has a wider range of application including camera stabilization, aircraft control, automobile and special switches. The inclinometer can be single axis device or dual-axis device for measuring roll and pitch. Inclinometer may be used to replace an accelerometer when the attitude is known (Kelly, 1996).

3.4 Odometry and differential odometry

Odometry is the estimation of the distance travelled by a wheeled vehicle relative to a starting point by use of information about the rotation of the wheel. In land vehicle navigation, dead reckoning is performed using several sensors combinations: odometer and gyroscope or odometer and compass or simply by differential odometry. Differential odometry extrapolates movement increments by differential encoders from which position and orientation can be estimated. Therefore differential odometry is only possible using a pair of encoder perpendicular to the direction of vehicle motion. There are two types of differential odometry,

synchronous drive in which each wheel is capable of being driven or steered and differential drive in which only rear or front wheels are driven or steered.

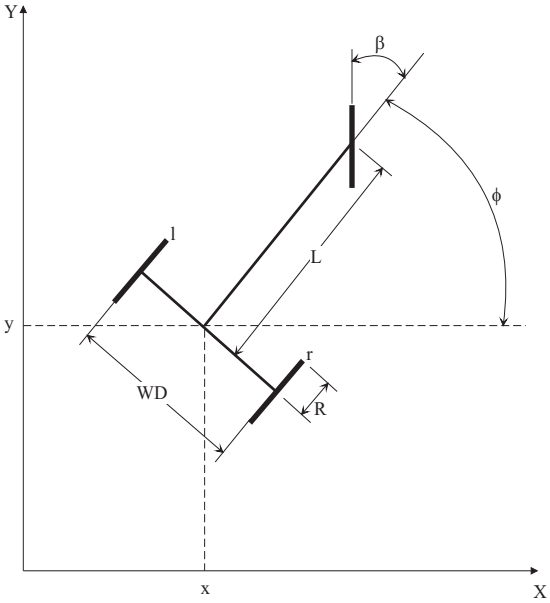

Figure 3.6: *Differential* o*dometry using wheel speed sensors*

Given a two wheel sensors, differential odometry estimates position and orientation from left and right wheel speed as a function of time. Differential odometry may also provide real-time position information in Cartesian (x and y) coordinates and orientation from periodic absolute measurements using shaft encoders. Figure (3.6) shows a differential odometry model and assuming that the vehicle moves in the plane and that the wheel velocities lie along the axis heading of each wheel then

$$\Delta\theta = \frac{(\Delta P_L - \Delta P_R)}{PPT \times WD} 2\pi R = \frac{(\Delta P_L - \Delta P_R)}{PPT \times WD}(360°) \times R \tag{3.8}$$

$$\Delta d = \frac{(\Delta P_L + \Delta P_R)}{PPT} \pi \times R \tag{3.9}$$

where $\Delta\theta$ is the heading difference, Δd is the distance difference, ΔP is the pulse difference, R is the radius of the vehicle wheel, PPT is the number of pulse signals generated per wheel turn and WD is the distance between the wheels (wheelbase). Taking x_i, y_i as the vehicle position and θ_i as the vehicle heading at an epoch i, the vehicle position at an epoch $i+1$ can then determined as shown in Equations (3.10) and (3.11)

$$y_{i+1} = y_i + \Delta d_{i+1} \sin(\theta_i + \Delta\theta_{i+1}) \tag{3.10}$$

$$x_{i+1} = x_i + \Delta d_{i+1} \cos(\theta_i + \Delta\theta_{i+1}) \tag{3.11}$$

3.4.1 Differential odometry errors

The wheel speed dead reckoning is subject to two main sources of error although errors resulting from numerical drift are also significant:

- Systematic errors: the systematic errors include imperfections in the manufacture such as unequal wheel diameters, wheel misalignments, uncertainty in the wheel base and limited encoder sampling rate and resolution. The systematic errors can be estimated and adjusted for through calibration.

- Non-systematic errors: these are due to skidding, slippage, thermal sensor noise and uneven ground. As a result of skidding the computed orientations can become totally unreliable.
- Numerical drift: due to discrete time integration of the displacements.

Road unevenness: the registration differential path difference as change in vehicle heading. Carlson *et al.* (2003) state that the primary limitation to differential odometry performance in heading estimation during normal driving is road surface unevenness and not wheel speed sensor resolution or wheel slippage.
Wheel slippage: the tires slip on wet, gravelled or icy roads. This should be mitigated.
Vehicle chassis: the wheel base and centre of gravity misalignments may cause errors in the computation of velocity and heading differences.
Scale factor determination: the scale factor i.e. a predetermined value fur distance per pulse signal count, relates the wheel displacements measured to the distance travelled. The scale factor should be determined using straight lie drives to avoid the influence of vehicle turnings. Zhao (1997) proposes the use of GPS to calibrate the ABS sensors scale factors when the vehicle is known to travel along a straight line.

3.4.2 Differential odometry resolution

Assuming a wheelbase of *1.6 m*, a signal per turn rate of *52* and a wheel radius of *33 cm* and based on *1* pulse per second error, i.e. worse case scenario (Hay, 2005), will result in the following error in heading differences

$$\varepsilon_{\Delta\theta} = \frac{1}{PPT \times WD} 2\pi R = \frac{1}{52 \times 1.6}\left(0.33\right)\left(360°\right) = 1.428° \tag{3.12}$$

As shown in Equation (3.12), the estimation of the vehicle heading differences using differential odometry is not useful for many applications due to the low resolution.

3.5 Integrated multi-sensor systems

Interests have grown in the use of multiple sensors to improve on the capabilities of the navigation and mapping sensors. The integration is usually performed to achieve redundancy and/or complimentary. The integrated system to use the information provided by the multiple sensors effectively, an integrating technique is required. Several integrating techniques have been suggested but Kalman filtering is still preferred because it produces results that are optimal in statistical sense. In land based positioning and navigation, we have two main categories of integrated multi sensor systems; position and orientation systems and mapping systems.

3.5.1 Position and orientation systems

The position and orientation systems are designed to measure the position and orientation of an object in dynamic environment. They are currently used in a variety of applications such as road profiling, geographical data acquisition and mapping. The main advantage of POS is that they provide accurate, continuous and consistent georeferencing data through integration of inertial measuring units and global positioning systems receivers.

Integrated INS/GPS systems

Latest development of integrated, extremely small, low cost GPS chips has opened new avenues by further improving location determination in navigation. Differential GPS has improved the position and orientation accuracy but GPS integration with other sensors and systems have done the "wonder". The reason for the integration of GPS and other sensors is to improve the positioning accuracy and reliability, a subject that has been vastly researched and extensively written upon. These include, for example, multipath effects in urban canyons, dynamic receiver limitations such as the jerk and signal outages in dense vegetation cover, undercover parking and in tunnels. GPS is a line of sight sensor technology which requires at least four satellites in view to solve a unique position. Hence for the GPS positions to be accurate and continuously available, the GPS positions should be integrated using other external sensors/aids. Some examples include GPS/INS, GPS/GIS (map-matching), GPS/LRF (laser range finder with built in electronic compass) and lately GPS/Cellular (mobile phones for emergency locations). After the accuracy and availability requirements are fulfilled, then the integrity of the system can be justified. Kinematic positioning using GPS has been applied for a while in precision navigation, surveying and mapping systems. The primary motivation for the integration of GPS and INS is the complementary advantages of the systems in overcoming each other's limitations (Babu and Wang, 2005). Titterton and Weston (2005) distinguish four main Kalman filter integration modes namely; *uncoupled, loosely coupled, tightly coupled* and *deeply (ultra-tight) coupled*.

Uncoupled GPS and INS integration mode simply uses the GPS position to reset the INS position at regular intervals of time. Loosely coupled integration considers the GPS and INS as independent systems and therefore integrates their position output externally in a Kalman filter. The main disadvantage of this mode is that there is no solution when the satellites are less than four. Tightly coupled GPS and INS integration mode combines the INS data with the GPS pseudo-ranges or carrier phases (Sennott and Senffner, 1997; Sennott, 1999). The main advantage of the tightly coupled Kalman filter integration is that the GPS and INS are treated as sensors making the solution possible even when the satellites are less than four. Ultra-tight GPS and INS integration mode integrates the GPS measurements and INS measurements by using the INS derived Doppler to remove the Doppler effects in the GPS signals through a feedback process. The deeply coupled Kalman integration mode has the advantage of higher GPS measurement accuracy by improving jamming to signal ratio and reducing RF signal interference.

3.5.2 Mobile mapping systems (MMS)

Mobile mapping system (MMS) is an information technology developed in the late 80s as a result of technological advances in positioning, communication, spatial information and other embedded technologies. All mobile mapping systems configuration usually consist of a moving platform on which the georeferencing sensors and mapping sensors are mounted, a synchronisation scheme, data integration and error processing software. The georeferencing sensors consist of GPS, DGPS, INS, dead reckoning, odometer, barometer, inclinometer, extensometers and laser range scanners while the mapping sensors include video cameras, charge couple devices (CCD) cameras and laser scanners. Sensors calibration and integration are also a part and parcel of mobile mapping technology. Automatic object recognition and extraction from mapping databases is critical. Figure (3.7) shows the main components of a MMS.

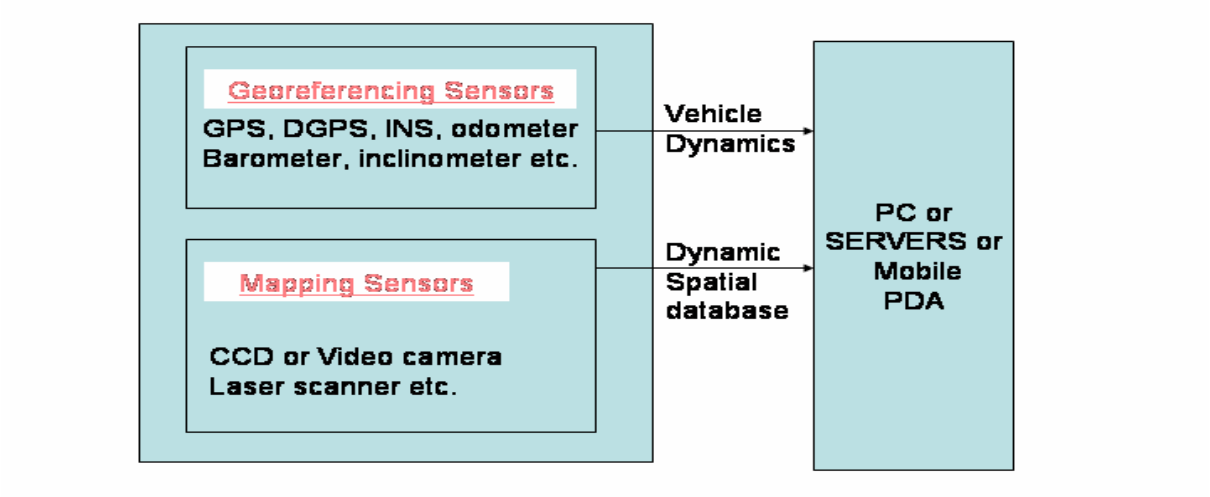

Figure 3.7: *Components of a MMS*

MMS is being applied in several fields including forestry and agriculture, intelligent transport, cadastral surveys and road inventory surveying, environment, engineering, planning, law enforcement, fire departments, construction inspectors, GIS oil industry and mining. Our interest is on the intelligent transport systems which include the so called land-based MMS.

Table 3.1: *Characteristics of sample MMS available today*

System		Position			Imaging		Other Sensors
Name	Developer	GPS	INS	DR	CCD	VHS	
GPS Van	OSU	√		√	2	2	
KISS	UBW	√	√		2	1	
LD2000-R	UW	√		√	2	1	
Visimind's MMS (Both land and Air)	Visimind AB	√	*		2		Orientation sensors
Portable MMS	UofC	√	*		1		Orientation sensors
VISAT	UofC	√	√	√	8		
Applanix POS LV™ system	Applanix						
GPSVision	Lambda Tech. Intern.	√	√		2		
ON-SIGHT	Transmap corp.	√	√		5		
GIGI-ONE	UofT						
MoSES	UFAF	√	√		2	1	Pointing laser
***Legend**: OSU-Ohio State University, UFAF-University of Federal Armed Forces, UBW-Univ. der Bundeswehr Munchen, UW-Wuhan University, UofC-University of California, UofT-Univ. of Trieste.*							

Most mobile mapping systems are engineered for specific purposes e.g. railway inventory survey. There is a growing number of MMS being developed in the world and some examples of the land based MMS developed for commercial and research purposes are listed in Table (3.1).

3.5.3 Applanix POS LV/AV systems

Applanix integration of inertial data with GPS has produced products for geomatic applications using sophisticated and advanced algorithms and signal processing techniques. Some of the systems from the Applanix Corporation developed for various applications include (http://www.applanix.com/); *POS AV*™ developed for direct georeferencing of airborne sensors data and performs geometric corrections and geocoding of the airborne data. Three models of *POS AV*™ are currently in the market i.e. *310, 410* and *510* in order of increasing quality. *POS LV*™ was developed for integrated position and orientation solution for the road vehicle-based applications while *POS TG*™ is for the computation of non-contact track geometry measurements and rail road surveying. *POS MV*™ is for high accuracy integrated multibeam sonar-inertial technology solutions applied in hydrographical surveying and lastly, *POS LS*™ is for land-based positioning.

Applanix POS LV™ *system*
The Applanix POS LV™ system is a tightly coupled inertial/GPS system as shown in Figure (3.8) and blends the INS data with GPS observations. The system offers the advantages of positioning with fewer satellites, reduces the re-acquisition time during RTK surveys and improved accuracy and integrity in the resulting navigation solution. POS LV™ components include a distance measuring instrument (DMI), an IMU and a GPS receiver.

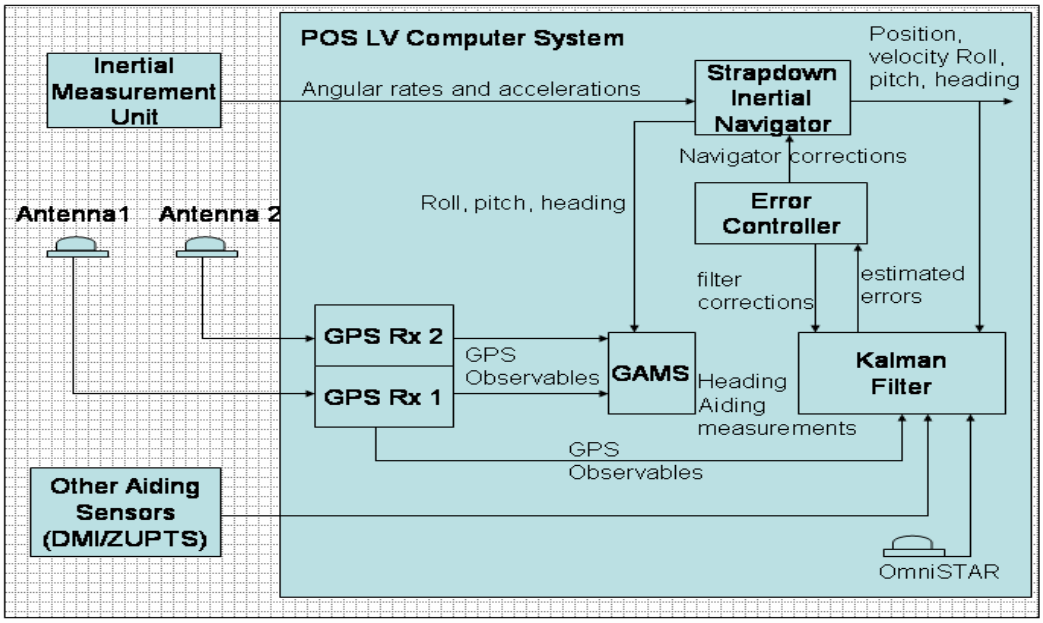

Figure 3.8: *POS LV Tightly Coupled System Architecture* (After Whittaker and Nastro, 2006)

Applanix POS AV™ *system*

The Applanix POS AV™ 410 is a hardware and software system, available at the institute of Navigation. The system is a six degrees of freedom strap-down integrated navigation system specifically engineered for aerial cameras, laser scanners, imaging sensors, synthetic aperture radar and LIDAR and consist of the following subsystems:

i. the IMU of type LN 200 consisting of a triad accelerometers, a triad gyroscopes and a digital/analogue converter
ii. GPS receiver for position and velocity measurements and also gives PPS for synchronisation purposes
iii. Personal computer for interfacing, processing and logging of data from the IMU and GPS receiver sub-systems.

The post processed data using POS Proc software gives integrated solution consisting of geographic position (latitude, longitude and altitude), velocity, acceleration, angular rate, orientation (attitude), heading and performance metrics for monitoring sensor health and for self configuration. The estimation process and system control is done through Kalman filtering.

4 Curve Fitting Functions

Equation Section 4

4.1 Introduction

A curve can be defined as a one dimensional connected set of points in a two dimensional or a three dimensional space. Curve fitting is the estimation of a curve which matches a series of data values and possibly other constraints. The need to perform a surface or a curve fitting process on a set of observed data points occur in many scientific disciplines involving numerical analysis such as CAD/CAM systems, quality control, engineering, image processing and finite element (FE) systems. Curve fitting is therefore done for different purposes or objectives, four of which are (Dierckx, 1995);

- *Parameter estimation*: a set of parameters defining a specific physical relation or a chosen function are estimated as accurately as possible from the given data points. This is also referred to as the *constructive approach* to curve fitting. For example, given data values y_i, for $i = 1, 2, ...n$ of the dependent variable y corresponding to the data values x_i, for $i = 1, 2, ...n$ of the independent variable x are fitted using a function $f(x) := y(x, \theta)$ of known form but containing the vector θ of m disposable parameters to be determined using an curve fitting criteria on the condition that $f(x_i) \approx y_i$.
- *Functional representation*: this is the derivation of a function that best approximates the data set and from which derivative values, integrals, etc. can be obtained. A mathematical function of specified form (for example a straight line, a polynomial or a spline) is then obtained through the curve fitting process by relating one dependent variable and one or more independent variables. The functional representation therefore involves the determination of the numerous functions that best fits a given data set and other constraints. This is also referred to as the *variational approach* to curve fitting in which, the functional form of $f(x)$ is not known but is determined in the solution of a variational problem.
- *Data smoothing*: the measurement errors in the data are smoothed out with the hope that the fitted curve is more accurate and smooth.
- *Data reduction*: this is a curve fitting process done with the aim that the estimated parameters should be less than the data values.

Our aim in this study is to perform a curve fitting process on a set of data points for the purposes of data smoothing and functional representation. All definitions and formulations in this chapter are based on a two dimensional space.

Given a sequence of distinct independent variables and for each a corresponding dependent variable, a pair of which is referred to as a *data point*, there are three ways of describing curves mathematically, *implicitly*, *explicitly* or as *parametric functions*. An implicit equation of a curve lying in the xy plane has the form $f(x, y) = 0$ thus describing the implicit relationship between x and y data points. On the other hand, a parametric function represents a point on the curve separately as a function of an independent parameter say u (*cf.* Equation (4.1)).

$$f(u) = (x(u), y(u)) \qquad a < u < b \tag{4.1}$$

where $f(u)$ is a vector-valued function of the independent variable u lying on the interval $[a\ b]$ normally normalised to $[0, 1]$. x and y are also called "measured" dependent variables. The widely used forms of curve representations in geometric modelling are *implicit* equations and *parametric* functions (Piegel and Wayne, 1997). Table (4.1) shows a summary of the different curve representation methods.

Table 4.1: *Examples of implicit, explicit and parametric functions*

implicit	explicit	parameterized
$h(x,y) = x^2 + y^2 - r^2 = 0$	$f(x) = \pm\sqrt{r^2 - x^2}$	$f(t) = \begin{pmatrix} r\cos t \\ r\sin t \end{pmatrix}$

4.1.1 Definition of the curve fitting terms

- *Control points*: Control points are data points in object space that affect the shape of the curve. The curve may pass near or through some of the control points. The control points are important to the representation of curves because they dictate the form of the parametric equations of the curve. Each control point therefore acts like a magnet tugging on the curve thus defining its shape.

- *Knot sequence*: The control points that lie on the curve are called knots. Knots define how the control points affect the curve. The more the number of knots the less the number of coefficients required to fit the curve. The knots determine whether the curve passes through and interpolates between some of the control points or passes near the control points.

- *Order*: The maximum multiplicity (maximum number of repeating a knot) is the *order* of the curve. The order of a curve affects the smoothness of the curve by influencing the shape of the curve when a control point is moved. Thus, the *order is equal to the number of knots minus the number of control points or one plus the maximum degree of the variable in the parametric equation*. The order is the number of coefficients in the parametric equation.

- *Continuity*: Duplicating values in the knot sequence increase that value's multiplicity but causes a decrease of the continuity of the curve. Continuity of the curve therefore depends on the order and the multiplicity i.e. *continuity = order - (multiplicity + 1).* Continuity determines the form of the parametric equations. No continuity means the curve is broken while zero-order continuity denoted C^0, means positional continuity between two curves i.e. the two curves are simply connected. First-order geometric continuity denoted C^1, means there exists an identical tangent between two curves or simply that the tangents to each curve at the point of connection have the same direction. Second-order curvature continuity denoted C^2 means that the two curves have the same curvature at the point of connection. Therefore, curvature continuity also implies both tangential and positional continuity.

- *Constraint*: refers to a point, angle or curvature that a curve fitting solution must fulfil. When the angle and curvature are added to the ends of the curve, as the case usually is, they are referred to as *end conditions*.

4.2 Curve fitting criteria

There are two types of curves used in curve fitting of data sets; *interpolating* curves and *approximating* curves. An interpolating curve passes exactly through the data points while an approximating curve passes as close and as smooth as possible to the data points. Thus depending on the type of the curve used, the curve fitting process is classified into two curve fitting criteria; curve interpolation or curve approximation criteria.

4.2.1 Curve interpolation criteria

Given a set of data points, curve interpolation involves finding a curve of a known degree that contains all the data points in the given order. Although interpolation is the simplest, it is less liable to unwanted fluctuations due to random errors in the data points. Interpolating curves are therefore the preferred approach in solutions where the data values are free of random noise otherwise approximating curves are applied. On the other hand, *global interpolation* estimates a function that fits all the data points while *piecewise interpolation* estimates a function that fits a portion of the data points.

4.2.2 Curve approximation criteria

Curve approximation criteria overcome the strict requirement of curve interpolation that the curve must contain all control points except the first and last points. Apart from the degree of the curve, the number of control points must also be known. There is no general rule in choosing the number of the knots, although the more the number of knots, the closer is the curve to the data points and when the number of knots m equals the degree of the function k, the approximation reduces to interpolation. Therefore, given a set of data points, curve approximation involves finding a curve of a known degree that contains the first and last data points and approximates the data points in the sense of certain curve approximation criteria. Some of the commonly used curve approximation criteria are least squares, natural smoothing and smoothing criteria.

4.2.2.1 The least-squares criterion

The least squares curve approximation criterion is based on the minimization of the sum of the squares of the error distances (residuals) between the data points and the curve fitting function. In the case of least-squares criteria, an approximated function $s(x)$ is determined under the condition that

$$e := \sum_{i=1}^{n} w_i (y_i - s(x_i))^2 = \text{minimal} \tag{4.2}$$

where w_i are the weights and e the closeness of fit (Dierckx, 1995 pp.45). The normal equations resulting from the least-squares process are solved using an appropriate method. There are two types of least squares approximations that can be performed. Fixed knot approximation is performed in the case where the data points, order and knots are given and only the coefficients are considered as unknown. Free knot approximation is performed when the data points and order are given while the coefficients and knots are considered as unknown.

4.2.2.2 The natural smoothing criterion

In the case of natural curve approximation criteria, a function $y(x)$ is determined such that

$$\eta := \int_{x_1}^{x_n} \left(y^{(m)}(x) \right)^2 dx = \text{minimal} \tag{4.3}$$

subject to the condition that

$$e := \sum_{i=1}^{n} (w_i (y_i - y(x_i)))^2 \leq \rho \tag{4.4}$$

where ρ is a specified non-negative number which acts as control (smoothness factor), m is the norm and η is a measure of non-smoothness of $y(x)$ (Dierckx, 1995 pp.46). Hence η should be as small as possible while e should be large enough. This is a constrained variational approach problem

The reason of performing smoothing curve approximation is mainly to remove the unwanted details and random noise in the data points. The parameter ρ can be estimated through one of the following methods

- can be specified as a degree of freedom
- can be estimated such that the cross validation (CV) is minimized. The concept of cross validation is using part of the data set for training (calibration) and the other part as test (validation) sample by switching the roles of the training and test samples.
- can be estimated such that the generalised cross validation (GCV) is minimized.

All in all, the choice of parameter ρ will depend on which of the following two conditions is given preference: the desire to stay close to the data points or the desire to obtain a smooth approximating function (or curve).

4.2.2.3 The smoothing criterion

This criterion tries to combine the advantages of the previous two criteria (Dierckx, 1975, 1981, 1982). These basic curve fitting criteria can be extended in different ways for example by imposing additional constraints. This involves approximating a curve of specified degree k, over a of knot interval $(\lambda_i, \lambda_{i+1})$ as solution of the following minimization problem. That is minimize (Dierckx, 1995, pp50)

$$\tilde{\eta} := \sum_{i=1}^{g} (s^{(k)}(\lambda_{i+1}) - s^{(k)}(\lambda_i))^2 \tag{4.5}$$

subject to the constraint

$$e := \sum_{i=1}^{n} (w_i (y_i - y(x_i)))^2 \leq \rho \tag{4.6}$$

where λ_i are the knots sequence for $i = 0, 1, \ldots, g+1$.

4.3 Types of curve fitting functions

The general requirements of the curve fitting function are: smoothness (i.e. no discontinuities in position, slope, direction and curvature), local control (i.e. the change of a data point should have only a local effect) and the curve should be intuitive and easy to use. In this section we examine three types of curve fitting functions in detail. These are straight lines, polynomials and splines.

4.3.1 Straight lines

The simplest forms of curve fitting functions are the lines where the relationship between the independent variable x and the dependent variable y is assumed to be linear as in Equation (4.7).

$$y = ax + b \tag{4.7}$$

Two methods can be used to perform a line fit to the data points; *slope and intercept* method or *linear regression* method. Slope and intercept method involves the determination of the slope of the line a and the intercept y i.e. the value when x is zero. In linear regression, the chi-square estimator $\chi^2(a,b)$ given in Equation (4.8) is estimated.

$$\chi^2(a,b) = \sum_{i=1}^{n}\left(\frac{y_i - b - ax_i}{\sigma_i}\right)^2 \tag{4.8}$$

where χ^2 is the function of the parameters of the model, σ_i are measurement errors set to 1 if not known and n is the total number of data points. Finding the values of a and b that minimize χ^2 involves differentiating the function with respect to a and b and setting the derivatives equal to zero (*cf.* Equations (4.9) and (4.10)).

$$\frac{\chi^2}{\partial a} = -2\sum_{i=1}^{n}\frac{x_i(y_i - b - ax)}{\sigma_i^2} = 0 \tag{4.9}$$

$$\frac{\chi^2}{\partial b} = -2\sum_{i=1}^{n}\frac{y_i - b - ax}{\sigma_i^2} = 0 \tag{4.10}$$

The resulting system of two equations can then be solved for a and b. Although linear interpolations are quick and easy, they are rarely precise and the resulting function is not differentiable at any data point.

4.3.2 Polynomial curves

A polynomial curve or a polynomial in short is a function that can be represented as a power series of the form

$$f(x) = b_0 + b_1x + b_2x^2 + \ldots\ldots + b_kx^k \tag{4.11}$$

or as Chebychev's series of the form

$$f(x) = b_0T_0 + b_1T_1(x) + b_2T_2(x) + \ldots\ldots + b_kT_k(x) \tag{4.12}$$

where $T_k(x)$ is the Chebyshev's polynomial of degree k (k being a none-negative integer), x is a scalar-valued variable and $b_0 \ldots\ldots b_k$ are fixed scalars called coefficients of the polynomial f (Cox and Hayes, 1973). Chebyshev's polynomial have two advantages as compared to other polynomials. First they allow for even distribution of errors over the whole approximation by maintaining the maxima error constant over the interval. They are numerically stable to compute since they can easily be truncated especially when polynomials of varied degrees are used in the assessment (Householder, 1953). The degree of the polynomial determines the number of coefficients used to approximate the polynomial. Polynomials can also be written in sigma form as

$$f(x) = \sum_{i=0}^{k} b_ix^i \tag{4.13}$$

Other possible representations of the polynomials are Newton form, Lagrange form, McLaurin series or Taylor series. Monomials, binomials and trinomials are other special forms of the polynomial with one, two or three terms respectively. Polynomials are simple and mathematically easily understood. They are easily, efficiently and accurately processed in computations and in particular with discrete points and their derivatives. The polynomials also provide good local approximates for smooth functions but on large intervals, the accuracy tends to be low. Moreover, local changes have global influence and are therefore very sensitive to choice of interpolation points. This global dependence on local properties can be avoided by use of piecewise polynomial approximants (De Boor, 1978; Höllig, 2001).

4.3.3 Splines

Spline curves were invented in the aircraft and ship building industries and were named after a loft man's flexible strip of material because they can be clamped or weighted to pass through any number of points with smooth deformation. Spline, a term coined by Schoenberg (1946), is a function whose sufficiently small first derivative approximates the shape of a draftsman's spline. Spline function consists of a series of polynomial functions over adjacent interval with continuous derivatives at the end points of the interval. Splines can, therefore, be simply defined as smooth piecewise parametric polynomials defined mathematically by two or more control points under certain conditions. In the design of splines one needs a set of knots, control points and coefficients.

Thus a function $s(x)$ defined on a finite interval $[a\ b]$ is called a spline of degree $k>0$ (order $k+1$) having as knots the strictly increasing sequence $t_i,\quad i=0.....n$ where $t_0=a$ and $t_n=b$ if the following conditions are satisfied:

i. On the interval $(t_i\ ,t_{i+1})$, $s(x)$ is given by a polynomial of degree k at most.

$$s|_{t_i,t_{i+1}}\in P_k \qquad i=0.....n \tag{4.14}$$

ii. The function $s(x)$ and its derivatives up to order $k-1$ are all continuous on the interval $[a\ b]$ i.e.

$$s(x)\in C^{k-1}[a,b] \tag{4.15}$$

From the second condition, one notes that the coefficients cannot be chosen freely and the degree k of the spline, the number and position of the knots and the coefficients in the representation determines the approximating function or criterion used (Dierckx, 1995). If the knot values are singular (no repeating values) and regularly spaced, the spline is a u*niform Spline*, otherwise it is referred to as *non-uniform Splines*.

Some of the commonly used types of splines are:

- Bézier curves
- B-splines
- Rational splines (Rational Bézier splines and Non-uniform rational B-splines (NURBS)).
- Cubic splines

The splines fall into two main categories. The first three spline categories are all *approximating splines* while the last type, the *cubic splines*, are either interpolating or interpolating splines. Somehow interpolation is the optimal fit but this is not advisable with the measured (real) data values due to measurement errors or noise. The best solution is to use other methods of approximation such as the smoothing criteria or least squares approximation

criteria. Splines have several advantages over the polynomials in that they are flexible and can easily be constrained. Cubic polynomials and splines are generally preferred in curve fitting since lower order functions have less flexibility while higher order functions are relatively complex to manipulate.

4.3.3.1 Bézier curves

A Bézier curve of degree d is defined over $n+1$ control points $P_0, P_1, \ldots\ldots P_n$ and a knot vector $U = \{u_1, u_2, \ldots\ldots, u_m\}$ as given Equation (4.16)

$$f(u) = \sum_{i=0}^{d} B_{i,d}(u) P_i \qquad 0 \le u \le 1 \tag{4.16}$$

where P_i are the control points and $B_{i,d}(u)$ are the classical Bernstein polynomial basis functions of degree d given by

$$B_{i,d}(u) = \binom{d}{i} u^i (1-u)^{d-i} \tag{4.17}$$

where $\binom{d}{i}$ is a binomial coefficient. Figure (4.1) below shows a Bézier curve defined over five control points. For a Bézier curve defined over four points, the basis function is given as shown in Equation (4.18).

$$\sum_{i=0}^{4} B_{i,3}(u) P_i = (1-u)^3 P_0 + 3u(1-u)^2 P_1 + 3u^2 (1-u) P_2 + u^3 P_3 \tag{4.18}$$

which in matrix form can be written as

$$\sum_{i=0}^{4} B_{i,3}(u) P_i = \begin{bmatrix} u^3 & u^2 & u & 1 \end{bmatrix} \begin{bmatrix} -1 & 3 & -3 & 1 \\ 3 & -6 & 3 & 0 \\ -3 & 3 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} P_0 \\ P_1 \\ P_2 \\ P_3 \end{bmatrix} \tag{4.19}$$

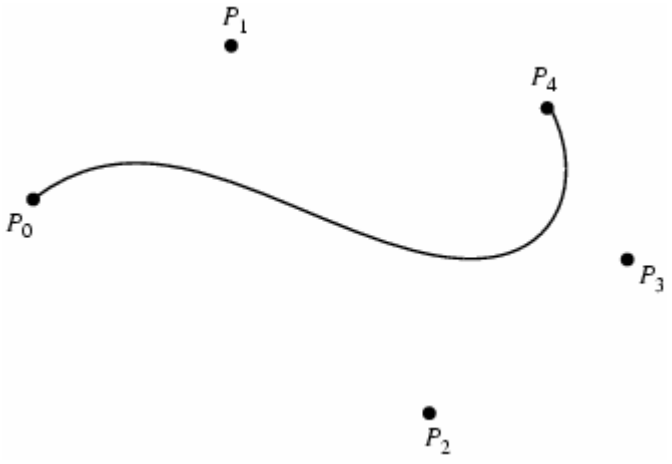

Figure 4.1: *A Bézier curve (*Weisstein, 2006a)

Therefore the point corresponding to independent variable u on the Bézier curve is the "weighted" average of all the control points where the weights are the functions $B_{i,n}(u)$. The geometric coefficients P_i are called control points. Note that $u \in [0, 1]$. As a result, all basis

functions are none-negative and passes through control points P_0 and P_n, also tangent to $P_1 - P_0$ and $P_n - P_{n-1}$ at the end points.

The shortcomings of the Bézier curves in geometric modelling of complex shapes are:

- that the require higher degrees to satisfy large number of constraints and model complex shapes e.g. $n-1$ degree is required to pass a polynomial Bézier curve through n data points. This results in numerical instability due to large number of control points
- that splitting the Bézier curve into piecewise segments represented by low order functions has control points which are not sufficiently local, that is, moving a single point changes the global shape of the spline

Bézier splines (piecewise Bézier curves) achieve first-order continuity by making the adjoining last three points collinear. Maintaining continuity requires several constraints on the control points with many curve segments. Therefore, although Bézier splines have local control property, they do not enforce derivative continuity at the join points.

4.3.3.2 B-spline curves

The basis splines or B-splines are defined on the fundamental theorem stating that every B-spline function of a given degree, smoothness and domain partition is a representation of a linear combination of B-spline basis functions of same degree and smoothness over that same partition. The B-spline basis functions can be defined using divided differences or through De Boor (1970) recursive relationship. B-splines are a generalisation of Bézier curves and involves more information namely $n+1$ control points $P_0, P_1, \ldots\ldots P_n$ and a knot vector of $m+1$ knots. Figure (4.2) shows a B-spline defined by six control point ($n=5$), and a sequence of ten knots ($m=9$). Take note of the knot multiplicity at the end of the B-splines, since for every given B-spline of degree d, the relationship of both the number of control points and the number of knots to the degree of the spline function $m = n + d + 1$ must hold.

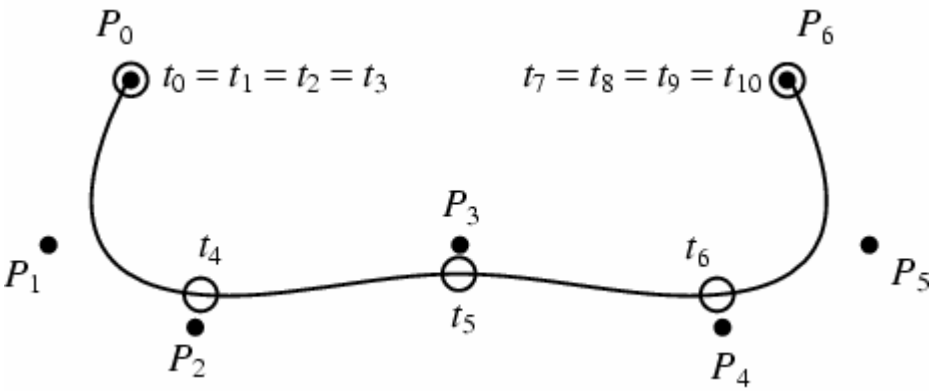

Figure 4.2: *A cubic B-spline curve with control points* (Weisstein, 2006b)

By applying the Leibniz's theorem, de Boor was able to derive the equation for the B-spline $C(u)$ as shown in Equation (4.20).

$$C(u) = \sum_{i=0}^{n} N_{i,d}(u) P_i \tag{4.20}$$

The underlying core being its basis or basis functions defined as

$$N_{i,0}(u) = \begin{cases} 1 & if \quad u_i \le u \le u_{i+1} \\ 0 & \text{otherwise} \end{cases}$$
$$N_{i,d}(u) = \frac{u - u_i}{u_{i+d} - u_i} N_{i,d-1}(u) + \frac{u_{i+d+1} - u}{u_{i+d+1} - u_{i+1}} N_{i+1,d-1}(u) \tag{4.21}$$

Equation (4.21) is the Cox-de Boor recursion formula where $N_{i,d}$ are the i^{th} B-spline basis functions of degree d or order $k = d + 1$, u_i are the non-decreasing set of real numbers also know as the knot sequence and u is the i^{th} knot span (interval).

A B-splines of degree d has C^{d-1} continuity at the common points between adjacent segments. $k+1$ knots or k intervals are required to define a basis function, each for every control point and the shape of the basis functions only depend on the knot spacing and not the specific knot values. Scaling or translating the knot vector does not alter the basis function. Similar to all other splines, when the knots are equidistant, we have a *uniform* B-spline otherwise it is called a *non-uniform* B-spline. For uniform B-splines, the basis functions are shifted copies of one another (*cf.* Figure 4.3).

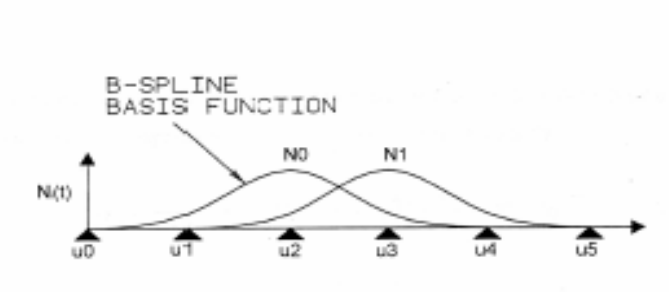

Figure 4.3: *Uniform cubic B-spline basis functions (After Farin, 1997)*

4.3.3.2.1 Computation of the B-spline basis functions

B-spline basis functions of degree one

The basis functions of zero degree (i.e. order $k = 1$), as per the first expression in Equation (4.21) are just step functions. For example for given four knots $u_0 = 0, u_1 = 1, u_2 = 2$ and $u_3 = 3$, knot spans *0, 1* and *2* are $[0,1], [1,2], [2,3]$ and the basis functions of degree 0 are $N_{0,0}(u) = 1$ on $[0,1]$ and 0 elsewhere, $N_{1,0}(u) = 1$ on $[1,2]$ and 0 elsewhere and, $N_{2,0}(u) = 1$ on $[2,3]$ and 0 elsewhere (*cf.* Figure 4.4) (Shene, 2006).

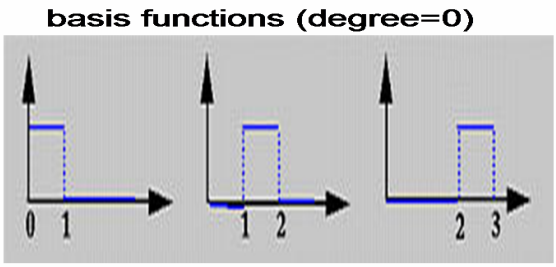

Figure 4.4: *B-spline basis functions of zero degree* (after Shene, 2006)

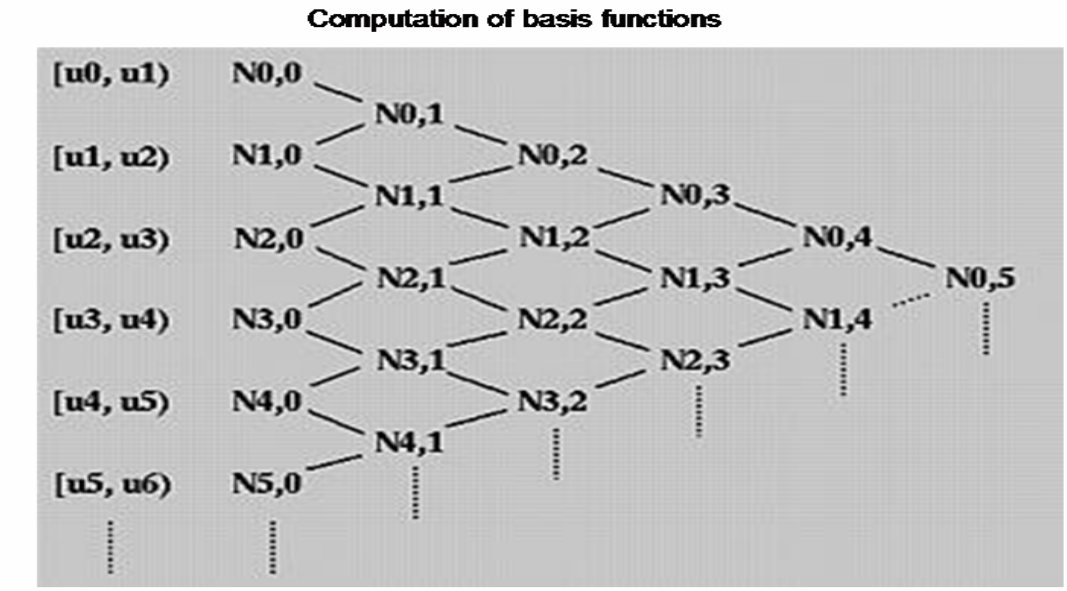

Figure 4.5: *Computation of B-spline basis functions* (after Shene, 2006)

B-spline basis functions of degree one (linear B-splines)

The computation of a basis function $N_{i,d}(u)$, of a degree d greater than zero is represented by the triangular scheme shown in Figure (4.5), where the knot spans are listed in the first column and all zero degree basis functions on the second column. In the computation of $N_{i,1}(u)$'s, the $N_{i,0}(u)$'s are required. Similarly to compute $N_{i,2}(u)$'s, the $N_{i,1}(u)$'s are required and so on. After substituting the previous values and rearranging, we have that

$$N_{0,1}(u) = \begin{cases} u & 0 < u < 1 \\ 2-u & 1 < u < 2 \\ 0 & \text{otherwise} \end{cases} \tag{4.22}$$

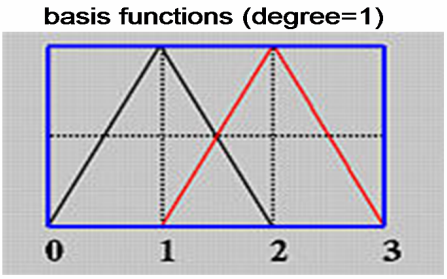

Figure 4.6: *Basis functions of degree one* (after Shene, 2006)

B-spline basis functions of degree two (quadratic B-splines)

Again in the computation of $N_{0,2}(u)$'s, the $N_{0,1}(u)$'s are required. Substituting the knot values in Equation (4.21) results in

$$N_{0,2}(u) = 0.5uN_{0,1}(u) + 0.5(3-u)N_{1,1}(u) \tag{4.23}$$

The values of the three knot spans are as given by Equation (4.24)

$$N_{0,2}(u)=\frac{1}{2}\begin{cases} u^2 & 0<u<1 \\ -2u^2+6u-3 & 1<u<2 \\ (3-u)^2 & 2<u<3 \\ 0 & \text{otherwise} \end{cases} \tag{4.24}$$

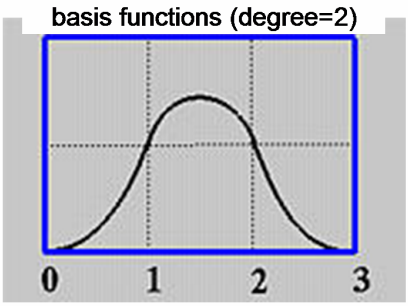

Figure 4.7*: Basis functions of degree two* (after Shene, 2006)

B-spline basis functions of degree three (cubic B-splines)

Similarly increasing the knot span by one, it can be shown that the B-spline basis functions of degree three are given by Equation (4.25). A detailed computation the basis functions are contained in Shene (2006).

$$N_{0,3}(u)=\frac{1}{6}\begin{cases} (u+3)^3 & 0<u<1 \\ -3u^3-15u^2-21u-5 & 1<u<2 \\ 3u^3+3u^2-3u+1 & 2<u<3 \\ (1-u)^3 & 3<u<4 \\ 0 & \text{otherwise} \end{cases} \tag{4.25}$$

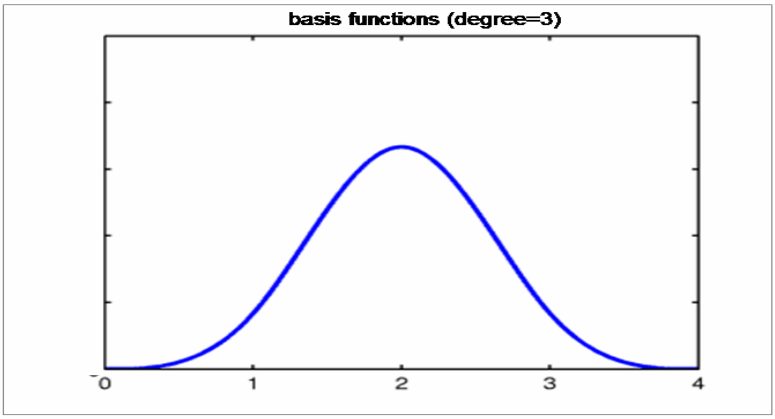

Figure 4.8*: Basis functions of degree three*

4.3.3.2.2 Curve fitting using B-spline basis functions

A constant B-spline is defined on only one knot and is not even continuous on the knots. A linear B-spline is defined on two knot spans and is continuous on the knots but not differentiable. A quadratic uniform B-spline is defined over three consecutive knots. Given a set of $n+1$ data points $D_0, D_1, \ldots, D_n$ and a set of parameters $u_0, u_1, \ldots, u_n$ corresponding to the data points, a desired B-spline curve of degree d is given as

$$D_k = C(u_k) = \sum_{i=0}^{n} N_{i,d}(u_k) P_i \qquad \text{for } 0 \le k \le n \tag{4.26}$$

Since $m = n + d + 1$, there are $n+1$ unknown control points and $n+1$ B-spline basis functions in the above equation. A matrix N of the basis functions in which the i^{th} row contains the values of $N_{0,d}(u), N_{1,d}(u), N_{2,d}(u), \ldots$ and $N_{i,d}(u)$, is computed at u_k are given in Equation (4.27)

$$N = \begin{bmatrix} N_{0,d}(u_0) & N_{1,d}(u_0) & N_{2,d}(u_0) & \cdots & N_{n,d}(u_0) \\ N_{0,d}(u_1) & N_{1,d}(u_1) & N_{2,d}(u_1) & \cdots & N_{n,d}(u_1) \\ \vdots & & & \ddots & \vdots \\ N_{0,d}(u_n) & N_{1,d}(u_n) & N_{2,d}(u_n) & \cdots & N_{n,d}(u_n) \end{bmatrix} \tag{4.27}$$

If the data points and the control points are given in two matrices D and P respectively in s-*dimensional* space

$$D = \begin{bmatrix} d_{0,1} & d_{0,2} & d_{0,3} & \cdots & d_{0,s} \\ d_{1,1} & d_{1,2} & d_{1,3} & \cdots & d_{1,s} \\ \vdots & & & \ddots & \vdots \\ d_{n,1} & d_{n,2} & d_{n,3} & \cdots & d_{n,s} \end{bmatrix} \text{ and } P = \begin{bmatrix} p_{0,1} & p_{0,2} & p_{0,3} & \cdots & p_{0,s} \\ p_{1,1} & p_{1,2} & p_{1,3} & \cdots & p_{1,s} \\ \vdots & & & \ddots & \vdots \\ p_{n,1} & p_{n,2} & p_{n,3} & \cdots & p_{n,s} \end{bmatrix} \tag{4.28}$$

then spline B-spline curve fitting boils down to the solution of a system of linear equations;

$$D_k = N \cdot P_i \tag{4.29}$$

where the D_k and P_i are the column vectors of D and P respectively. This is the process of global interpolation.

For B-spline curve approximation, two more inputs are required, that is, the number of control points $n+1$ and a degree of the spline d, where $m = n + d + 1$ must also hold. The approximation is therefore more flexible in that one only selects the degree but also the number of control points. With the same data points, Equation (4.26) is rewritten as in Equation (4.30).

$$C(u_k) = N_{0,d}(u) D_0 + \left(\sum_{i=1}^{h-1} N_{i,d}(u) P_i \right) + N_{h,d}(u) D_n \tag{4.30}$$

where $P_0, P_1, \ldots, P_h$ are the $h+1$ unknown control points and since the curve has to pass through the first and last data points then $D_0 = C(0) = P_0$ *and* $D_n = C(1) = P_h$. This is then solved, for example, using least squares criteria minimizing the squares of the errors between

the approximated curve point u_k and the corresponding data point D_k of on the curve $|D_k - C(u_k)|^2$. The control points that minimize the function

$$f(P,\ldots,P_{h-1}) = \sum_{k=1}^{n-1} |D_k - C(u_k)|^2 \tag{4.31}$$

are determined. This is global approximation process.

The unique characteristics of the B-spline curves are that

- they are positive within the interval $(0, n+1)$ and vanishes outside this interval
- they are $k-1$ times continuously differentiable with discontinuity of the k^{th} derivative
- they are piecewise polynomials of degree k on each interval $(i, i+1), k = 0.....k$

These conditions are valid so long as the control points are uniformly spaced and the polynomial segments join smoothly.

4.3.3.3 Rational splines

Non-rational splines (Bézier and B-splines) are polynomial parametric curves and can only be used to represent polynomials of the parametric form. They can not represent conic sections like circles and ellipses. To represent a conic section (*cf.* Figure 4.9), one needs additional degrees of freedom which can only be achieved using rational curves. In rational splines, each control point has an associated weight that influences the shape of its basis function. For the non-rational splines discussed above, all control points have a weight of one. The parametric equations for rational curves have both a numerator and a denominator, which results in a ratio. Each point is a ratio of two curves.

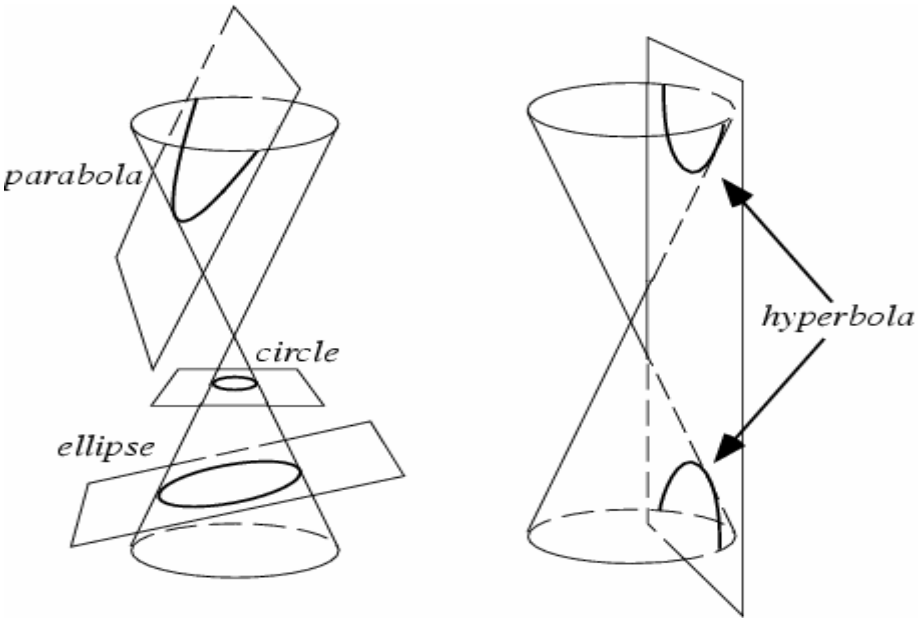

Figure 4.9: *Conic sections: circle, ellipse, parabola and hyperbola* (Weisstein, 2006d)

A rational curve can be seen as a quotient of polynomials. Bézier and B-splines are generalised to rational Bézier and Non-Uniform Rational B-splines (NURBS).

4.3.3.3.1 Rational Bézier curves

Given $n+1$ control points each of which is associated with non-negative weights w_i, a rational Bézier spline is given by;

$$f(u) = \sum_{i=0}^{n} R_{i,n}(u) P_i \tag{4.32}$$

Where $R_{i,n}(u)$ are the rational basis functions for the curve defined as (After Piegl, 1997).

$$R_{i,n}(u) = \frac{B_{i,n}(u) w_i}{\sum_{j=0}^{n} B_{j,n}(u) w_i} \tag{4.33}$$

Rational Bézier curves are special case of NURBS and therefore satisfy some of the properties of NURBS. They lack local control and therefore moving or changing a control point will have global effect.

4.3.3.3.2 NURBS

NURBS is a synonym for non-uniform rational B-spline which is a rational form of the B-spline curve and has the added advantage that it's invariant under perspective transformation. The terms non-uniform refers to type of knot vector while rational refers to the type of function. NURBS are defined by

$$f^w(u) = \sum_{i=0}^{n} N_{i,p}(u) P_i^w = \frac{\sum_{i=0}^{n} N_{i,p}(u) w_i P_i}{\sum_{i=0}^{n} N_{i,p}(u) w_i} \tag{4.34}$$

where $N_{i,p}(u)$ are the B-spline basis functions, P_i are the control points and w_i are the weights of P_i. The control points of the spline are multiplied by a non-zero number which does not change their values i.e.

$$P_i^w = \begin{bmatrix} wy_i \\ wx_i \end{bmatrix} \tag{4.35}$$

The weight is generally positive. If weight is zero, the control points are disabled while negative effects have other interesting applications. If the weight is 1, the NURBS reduces to a B-spline curve.

Compared to B-splines, NURBS have more degrees of freedom. In comparison to Bézier curves, NURBS have C^{k-1} continuity. The disadvantage of NURBS is that they are linearly dependent on the dimension but not the weights and therefore one cannot build a system of linear equations making them inappropriate for approximation and interpolation tasks. Circles, ellipses and other complex curves are possible with NURBS and have therefore become a standard form of representing complex geometric information in computer aided design (CAD), CAE and CAM. NURBS are also an integral part of many data exchange formats such as initial graphics exchange specifications (IGES), standard for the exchange of

product model data (STEP) and programmer's hierarchical interactive graphics system (PHIGS).

Table 4.2: *Summary of the characteristics of the splines*

	Effects of Control	Knot intervals	Blending function	other
B-spline	local	uniform	Normalised polynomial	Smooth
Bézier curve	non-local	uniform	Berstein polynomial	Last 3 control points must be linear for position continuity
NURBS	local	unrestricted	Rationalised polynomial	Linear dependence on dimension
Rational Bézier	non-local	unrestricted	Rationalised Berstein polynomial	

In summary, B-splines are generalisation of the Bézier splines i.e. a B-spline with no internal knots is a Bézier curve. Similarly, a Bézier spline is therefore a special case of B-spline. Figure (4.10) shows the relation between Bézier splines, B-spline, rational Bézier splines and NURBS.

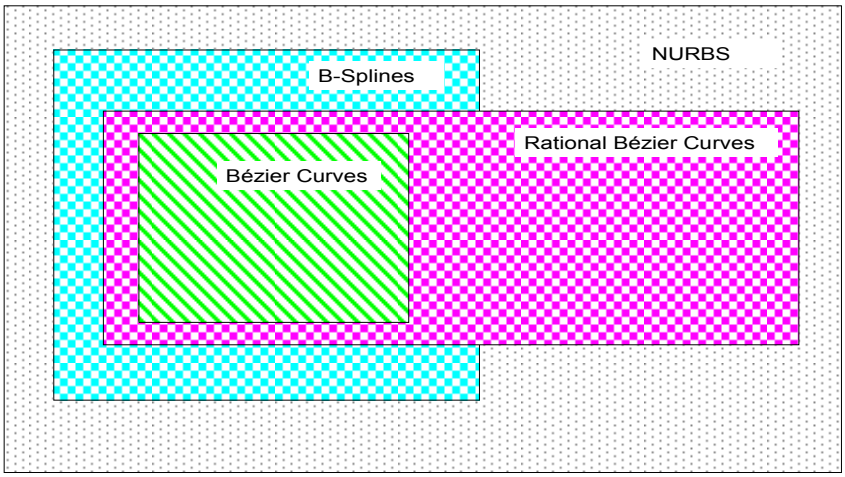

Figure 4.10: *Generalisation of the B-splines and Bézier curves (Redrawn after WIKIPEDIA)*

Since Bézier curves have the disadvantage that they lack local control, while NURBS are difficult to formularise into a set of linear equations making their solution by least-squares impossible. The alternative is therefore to apply least-squares criteria on B-splines to approximate the designed curves constituting the road geometry.

4.3.3.4 Interpolating Cubic splines

McKinley and Levine (1995) noted that *"..... a function which would effectively correlate the data would be difficult to obtain and highly unwieldy. To this end, the idea of cubic splines was developed. ....."*. Cubic splines consist of series of cubic polynomials filled between each of the data points with the condition that the interpolated spline is continuous and smooth. The algorithms for interpolation are classified into three methods; *projection method*, method of *basis function* and the *direct solution*.

According to de Boor (1978) p.49-50, given a set of data points $y(t_1), y(t_2), \ldots\ldots y(t_n)$ with $a = t_1 < \ldots < t_n = b$, a piecewise cubic spline $f(x)$ is constructed to the data set y_i on each interval (t_i, t_{i+1}). The function $f(x)$ must agree with a polynomial P_i of order 4.

$$f(t_i) = P_i(t_i) \quad \text{for} \quad t_i \le x \le t_{i+1} \tag{4.36}$$

The i^{th} polynomial piece P_i must satisfy the following conditions

$$\left.\begin{array}{ll} P_i(t_i) = y(t_i) & P_i(t_{i+1}) = y(t_{i+1}) \\ P_i'(t_i) = s_i & P_i'(t_{i+1}) = s_{i+1} \end{array}\right\} i = 1 \ldots n-1 \tag{4.37}$$

where $s_1 \ldots s_n$ are free parameters ("slopes"). The cubic function derived from y_i at $t_1 \ldots t_n$ are continuous and have continuous first derivatives on $[a, b]$ regardless of the choice of the free parameters s_i. In the computation of the coefficients of the i^{th} polynomial, the Newton form of polynomial representation shown in Equation (4.38) can be used. The coefficients are then derived from the divided differences as indicated in Table (4.3).

$$P_i(x) = P_i(t_i) + (x - t_i)[t_i, t_i]P_i + (x - t_i)^2[t_i, t_i, t_{i+1}]P_i + (x - t_i)^2(x - t_{i+1})[t_i, t_i, t_{i+1}, t_{i+1}]P_i \tag{4.38}$$

Table 4.3: *Divided differences (after de Boor, 1978: p.50)*

<table>
<tr><th></th><th>$y_i = [\ ]P_i$</th><th>$[,]P$</th><th>$[,,]P$</th><th>$[,,,]P$</th></tr>
<tr><td>t_i</td><td>y_i</td><td rowspan="2">s_i</td><td rowspan="2">$\left[\frac{y_{i+1} - y_i}{t_{i+1} - t_i} - s_i\right] / \Delta t_i$</td><td rowspan="3">$\left[s_{i+1} + s_i - 2\frac{y_{i+1} - y_i}{t_{i+1} - t_i}\right] / (\Delta t_i)^2$</td></tr>
<tr><td>t_i</td><td>y_i</td></tr>
<tr><td>t_{i+1}</td><td>y_{i+1}</td><td rowspan="2">$\frac{y_{i+1} - y_i}{t_{i+1} - t_i}$</td><td rowspan="2">$\left[s_{i+1} - \frac{y_{i+1} - y_i}{t_{i+1} - t_i}\right] / \Delta t_i$</td></tr>
<tr><td>t_{i+1}</td><td>y_{i+1}</td><td>s_{i+1}</td></tr>
</table>

The free parameters (slopes) s_i are a result of oscillatory interpolation process since there is a higher than first order contact between the function and its interpolant i.e. apart from the functional values, the derivatives are also interpolated. In terms of shifted powers $(x - t_i)^r$, the representation in Equation (4.38) can be rewritten as in Equation (4.39).

$$P_i(x) = C_{1,i} + C_{2,i}(x - t_i) + C_{3,i}(x - t_i)^2 + C_{4,i}(x - t_i)^3 \tag{4.39}$$

where

$$\begin{aligned} C_{1,i} &= P_i(t_i) = y_i \\ C_{2,i} &= [t_i, t_i]P_i = P_i'(t_i) = s_i \\ C_{3,i} &= [t_i, t_i, t_{i+1}]P_i = \frac{P_i''(t_i)}{2} = \left[\frac{y_{i+1} - y_i}{t_{i+1} - t_i} - s_i\right] \Big/ \Delta t_i - C_{4,i}\Delta t_i \\ C_{4,i} &= \frac{(x - t_{i+1})}{(x - t_i)}[t_i, t_i, t_{i+1}, t_{i+1}]P_i = \frac{P_i'''(t_i)}{6} = \left[s_i + s_{i+1} - 2\frac{y_{i+1} - y_i}{t_{i+1} - t_i}\right] \Big/ (\Delta t_i)^2 \end{aligned} \tag{4.40}$$

The construction of cubic spline is based on following principles;

- curves are third order polynomials

$$P_i(t) = a_i + b_i t + c_i t^2 + d_i t^3 \tag{4.41}$$

- curves pass through all known points (interpolating)

$$P_i(t_i) = P_{i+1}(t_i) = y_i \tag{4.42}$$

- The slopes or first order derivatives, are the same for both functions on either side of a control point

$$P_i(t_i) = P'_{i+1}(t_i) \tag{4.43}$$

- The second order derivatives, are the same for both functions on either side of a control point

$$s''_i(t_i) = s''_{i+1}(t_i) \tag{4.44}$$

The criteria are achieved by determining the free parameters (slopes) $s_2 \ldots s_{n-1}$ through the condition

$$s_{i-1}\Delta t_i + s_i 2(\Delta t_{i-1} + \Delta t_i) + s_{i+1}\Delta t_{i-1} = b_i \tag{4.45}$$

where

$$b_i = 3\left(\Delta t_i \frac{y_i - y_{i-1}}{t_i - t_{i-1}} + \Delta t_{i-1} \frac{y_{i+1} - y_i}{t_{i+1} - t_i} \right) \tag{4.46}$$

for $i = 2 \ldots\ldots n-1$

This results in $n-1$ equations with $n-1$ unknowns. The solution can be obtained by solving the tridiagonal system of linear equations using, for example, Gauss elimination without pivoting (de Boor, 1978). The remaining free parameters s_1 and s_n are obtained by imposing boundary conditions as in Section (4.3.3.4.1) below. The choice of the slopes s_i in Table (4.3) determines the type of the cubic spline interpolation (see Section 4.3.3.4.2).

4.3.3.4.1 Boundary Conditions

The quality of a cubic interpolating spline is somewhat inferior towards the end points of the approximating interval. This is due to lack of enough information i.e. a cubic interpolating spline requires four points to compute the coefficients of a data point. This situation can be solved partly by imposing boundary conditions (constraints). Depending on the end condition constraint applied, the different types of interpolating cubic polynomials are categorised as follows

Clamped or complete: sets the first derivatives at the start point and end point equal to given values normally first derivative at t_1 and t_n hence matches the end slopes to the specified values

$$s_1 = y'(t_1) \quad \text{and} \quad s_n = y'(t_n) \tag{4.47}$$

Lagrange: matches the end slopes to the slopes of the cubic curve that matches the first four points
As Lagrange polynomials oscillate between its roots (knots), therefore they can take negative values

Natural: sets the second derivatives at the start point and end point equal to zero i.e.

$$s_1 = 0 \quad and \quad s_n = 0 \tag{4.48}$$

Not-a-knot: first two segment and last two segments interpolate as single cubic curves i.e. making second and second last sites inactive knots. The result is that the third derivatives of the piecewise cubic polynomials are continuous. These are also known as *cubic runout splines* since the curve degrade to a single cubic spline over the last two intervals by setting the second order derivatives of the splines at the end points to

$$s_1''(t_0) = 2s_1''(t_1) - s_2''(t_2) \tag{4.49}$$

$$s_n''(t_n) = 2s_n''(t_{n-1}) - s_{n-1}''(t_{n-2}) \tag{4.50}$$

Variational: sets end second derivatives equal to zero

$$s_n'' = 0 \tag{4.51}$$

Periodic: matches the first and second derivatives at the left with those at the right i.e. the first and second derivatives at the left and right are equal

$$s_1 = s_n, \quad s_1' = s_n \quad \text{and} \quad s_1'' = s_n'' \tag{4.52}$$

Bessel's end condition: requires passing a quadratic parabola through the first three interpolation points.

The use of coefficients derived from cubic spline interpolation is to compute first and second order derivatives as well as curvature at the discrete data points. The curvature values are then analysed and used to determine the start and end points of the individual geometric road design elements i.e. tangents, circular curves and transition curves.

4.3.3.4.2 Types of Cubic Interpolation

Cubic Hermite Interpolation

The choice o the slope is made such that $s_i = y'(t_i)$ for all i thus making the approximation local. Cubic Hermite splines are third degree functions defined by two control points and two control tangents i.e. starting control point and tangent and ending control point and tangent. Several methods exist for estimating plausible tangent values where each two polynomial pieces meet and this have resulted into further classes of cubic Hermite splines such as cardinal spline, Overhauser (Catmull-Rom) spline, β splines and Kochanek-Bartels splines also called Tension Continuity Bias (TCB) splines

Cubic Bessel Interpolation

One chooses the slope parameter s_i as the slope at t_i of the polynomial of order 3 (degree 2) which agrees with y_i at t_{i-1}, t_i, t_{i+1} i.e.

$$s_i = \frac{\Delta t_i \dfrac{y_i - y_{i-1}}{t_i - t_{i-1}} + \Delta t_{i-1} \dfrac{y_{i+1} - y_i}{t_{i+1} - t_i}}{\Delta t_i + \Delta t_{i-1}} \tag{4.53}$$

hence the effects of changing the control points are local.

Akima's Interpolation

Akima's choice is

$$s_i = \frac{w_{i+1}\dfrac{y_i - y_{i-1}}{t_i - t_{i-1}} + w_{i-1}\dfrac{y_{i+1} - y_i}{t_{i+1} - t_i}}{w_{i+1} + w_{i-1}} \tag{4.54}$$

where

$$w_i := \left| \frac{y_{i+1} - y_i}{t_{i+1} - t_i} - \frac{y_i - y_{i-1}}{t_i - t_{i-1}} \right| \tag{4.55}$$

In Akima's choice, the effects of changing the control points are local.

4.3.4 Other functions

Some other curve fitting functions are pade approximants, trigonometric functions, wavelets and discrete Fourier transforms which are special trigonometric functions. Others are multivariate regression functions in which data sets with two or more independent variables e.g multivariate linear, multivariate polynomials are employed.

4.4 Differential geometry of the curve

Differential calculus invented by Leibniz and Newton solves the problem of finding the tangent to any given curve. Conversely one can obtain the equation of the curve by solving a differential equation from the relation between the curve and its tangent. As mentioned in Section (4.1), the implicit equation of a curve is given as $f\left(x(u), y(u)\right) = c$ where f is a function of x and y while c is a constant. The first derivative f' of a parameterised curve is called the tangent vector of f at the point $\left(x(u), y(u)\right)$. The arc length of the curve f from a point $\left(x(u_0), y(u_0)\right)$ to the point $\left(x(u), y(u)\right)$ is the function $s(u)$ given by

$$s(t) = \int_{u_0}^{u} \left\| f'(u) \right\| du \tag{4.56}$$

Curvature measures the extent to which a curve is not contained in a straight line (i.e. straight lines or tangents have zero curvature) or simply the amount by which a geometric object deviates from being "flat" plane. Torsion, on the other hand, measures the extent to which a curve is not contained in a plane (i.e. plane curves have zero torsion). The curvature and torsion describes completely the shape of a curve in space. In plane curves, the curvature of a curve at every point describes the shape of the curve completely. Although the curvature determines the shape of a curve it does not tell anything about the position of the curve.

The simplest form of curvature and that usually first encountered in calculus is an extrinsic curvature (Weisstein, 2004c). If a plane curve is given by parametric equations $x = x(u)$ and $y = y(u)$, then the curvature κ is defined by

$$\kappa = \frac{d\phi}{ds} = \frac{d\phi}{du}\frac{du}{ds} = \frac{d\phi}{du}\frac{1}{\sqrt{\left(dx/du\right)^2 + \left(dy/du\right)^2}} = \frac{d\phi}{du}\frac{1}{\sqrt{x'^2 + y'^2}} \tag{4.57}$$

where ϕ is the tangential angle and s is the arc-length. The assumption here is that the curve is a regular parameterised curve such that $f' = d/du \neq 0$ holds for every point. Curvature at any point is inversely proportional to radius of the curve at that point and has units of inverse distance. The $d\phi/du$ derivative in Equation (4.57) above can be found using the identity

$$\tan\phi = \frac{dy}{dx} = \frac{dy}{du}\frac{du}{dx} = \frac{y'}{x'} \tag{4.58}$$

Therefore

$$\frac{d}{du}\left(\tan\phi\right) = \sec^2\phi\frac{d\phi}{du} = \frac{x'y'' - y'x''}{x'^2} \tag{4.59}$$

and

$$\frac{d\phi}{du} = \frac{1}{\sec^2\phi}\frac{d}{du}\left(\tan\phi\right) = \frac{1}{1+\tan^2\phi}\frac{x'y'' - y'x''}{x'^2} = \frac{x'y'' - y'x''}{x'^2 + y'^2} \tag{4.60}$$

Combining Equations (4.57) and (4.60) gives

$$\kappa = \frac{x'y'' - y'x''}{(x'^2 + y'^2)^{3/2}} \tag{4.61}$$

For a two-dimensional curve in the form $y = f(x)$ (i.e. a graph), the equation of the curvature becomes

$$\kappa = \frac{d^2y}{dx^2}\frac{1}{\left[1+\left(dy/dx\right)^2\right]^{3/2}} = \frac{y''}{\left[1+y'^2\right]^{3/2}} \tag{4.62}$$

A regular curve has constant curvature κ if and only if, it is part of circular curve of radius $1/|\kappa|$ (if $\kappa \neq 0$) or a line segment (if $\kappa = 0$) (Gray 1997). The clothoids as discussed in Section (2.2.1.3) are spiral curves whose curvatures are linear functions of their arc-length. They are used in road design to smoothly link curves of infinite radius of curvature (straight tangents) to curves of finite radius of curvature (circular curves). Clothoids can be expressed parametrically in terms of Fresnel's integrals as

$$x = aC\left(u\right) \qquad y = aS\left(u\right) \tag{4.63}$$

where a is the scaling function > 0, u is a parameter that varies from $0 \to \infty$ while

$$C\left(u\right) = \int_0^u \cos\left(\frac{\pi u^2}{2}\right)du \tag{4.64}$$

$$S\left(u\right) = \int_0^u \sin\left(\frac{\pi u^2}{2}\right)du \tag{4.65}$$

are the integrals and the tangent angle at u is $\pi u^2/2$, arc-length $ds = adu$ and the curvature is $k = \pi u/2$, As $u \to \infty$, the clothoid plot approaches $\left(a/2, a/2\right)$ (Joy, 1999). The geometry of

the clothoid is interesting since it can be reconstructed simply from its curvature although the mathematics is somewhat complicated.

4.5 Curve parameterization of the measured road coordinates

Due to dimensional, analytical and combinatorial complexities, the geometric objects in plane are simpler as compared to those in spaces. The algorithm therefore splits the road geometry into two planes; the horizontal plane and the vertical plane. As mentioned in Chapter two, road designs typically involve the smooth construction of three fundamental elements: straights, circular curves and transition curves (spirals or clothoids). Curve parameterization has been defined by Armstrong (2006) as the specification of the parameter type and parameter values at interpolation points. The Cubic spline, as discussed in the previous section, is defined by a local parameter u that varies from *0 to 1* at sufficiently small increments while the entire curve is defined by another parameter t also varying from *0 to 1* at sufficiently small increments. Thus the global parameter t is mapped onto the local parameter u at each segment of the piecewise curve fitting process and are both determined at each knot. Knot spacing though plays no roll in the parameterization i.e. do not affect the shape of the curve but only affects the distribution of the markers along the curve.

The measured vehicle track is defined by a parametric curve whose curve parameter is time i.e. the x and y coordinates are defined in time series as $\left(x_{t_1}, y_{t_1}\right), \left(x_{t_2}, y_{t_2}\right), \cdots, \left(x_{t_n}, y_{t_n}\right)$, n being the number of the measured point coordinates. A parametric representation of the vehicle track is therefore given as in Equation (4.66)

$$f(t) = \left(x(t), y(t)\right) \tag{4.66}$$

Due to the varying speed of the vehicle during measurement, the time series result in non-uniform coordinate spacing. Therefore, the desired spline is evaluated at points based on the arc-length. The arc-length parameterization of the curve (spline) involves several steps (Gil and Keren, 1997; Wang *et al.*, 2002):

- estimate the arc-length of all the cubic spline segments of the input spline $f(t)$ as a function of the independent variable (parameter) in this case time t: $L = D(t)$, where L is a strictly increasing function of t and therefore have one-to-one correspondence.
- estimate equally spaced $m+1$ points along the spline $f(t)$.
- estimate $\tilde{t} = D^{-1}(L)$ corresponding to the equally spaced points.

Substituting $\tilde{t} = D^{-1}(L)$ into Equation (4.66) above, then the arc-length parameterized curve is given by Equation (4.67).

$$f(L) = \left(x\left(D^{-1}L\right), y\left(D^{-1}L\right)\right) \tag{4.67}$$

Several methods of evaluating $D^{-1}(L)$ have been suggested an example of which are the direct solution (Hartley and Judd, 1980), the root finding (Sharpe and Thorne, 1982), the cubic spline interpolation (Armstrong, 2006) and the Chebyshev approximation (Peterson, 2006). In this project, the bisection method described by Wang *et al.* (2002) was adopted as detailed in the process shown in Figure (4.11). The arc-length of each spline segment (i.e. the chord lengths) are approximated by

$$\Delta d \approx \sqrt{\left(x_{i+1} - x_i\right)^2 + \left(y_{i+1} - y_i\right)^2} \approx \sqrt{\left(\Delta x_i^2 + \Delta y_i^2\right)} \tag{4.68}$$

where i varies from 0 to $n-1$ and n is the number of spline segments in the original curve. Therefore $L(t)=\sum_{i=1}^{n-1}\Delta d_i$. Using bisection method, equally spaced points along the spline $f(t)$ are estimated such that $\Delta d_m = L(t)/m$ is the length of each segment and

$$L(t)=\sum_{i=1}^{m}\Delta d_i = m\cdot\Delta d_m \tag{4.69}$$

The parameterization starts with the search of the time interval $[t_i, t_{i+1}]$ containing the desired arc-length $L(m)$ such that $t_i \le \tilde{t} \le t_{i+1}$. The parameters $L(m)$ and $\tilde{t}$ are therefore solved iteratively using Equations (4.70) and (4.71) below.

$$L(m)=\sum_{i=1}^{m}\Delta d_i + \frac{\tilde{t}-t_i}{t_{i+1}-t_i}(\Delta d_m) \tag{4.70}$$

$$\tilde{t}=t_i+\frac{L(m)-L(t_i)}{L(t_{i+1})-L(t_i)}(t_{i+1}-t_i) \tag{4.71}$$

The parameterize spline curve is therefore estimated as given by Equation (4.72) .

$$f(L(m))=\left(x(L(m)), y(L(m))\right) \tag{4.72}$$

The fitted curve is then evaluated using standard deviation of residuals, goodness of fit and standard deviation of parameters.

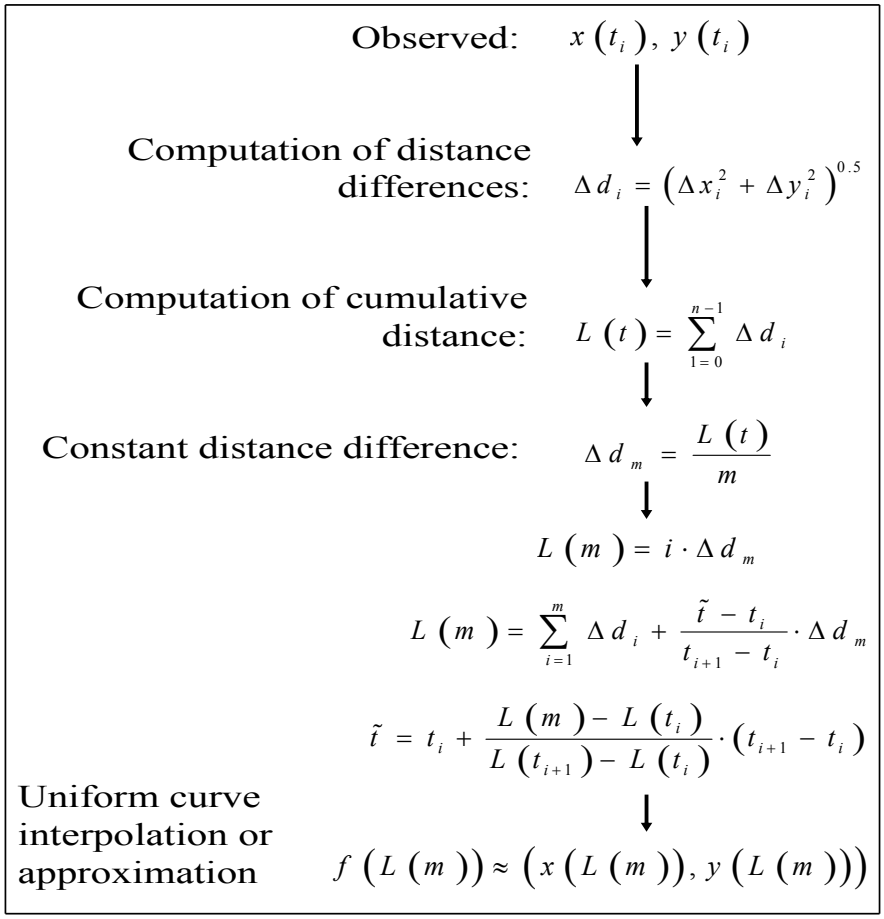

Figure 4.11: *Arc length parameterization of the measured data points*

5 Analysis of Error Effects on Geometric Road Design Elements

Equation Section 5

5.1 Introduction

As stated in Chapter (1), one of the aims of this study is to recover the geometric road design elements using a low-cost sensors. The design elements that describe the geometry of the horizontal road profile can be considered to be comprised of the straight lines or tangents, circular arcs and clothoids or transition curves. When high accuracy positioning sensors are used in the measurement of the road trajectory, there should be no or negligible discontinuities in position, direction and curvature between the individual geometric design parameters. The observed trajectory positions using low-cost sensors, however, contain stochastic and deterministic errors which must be reduced through filtering processes before the trajectory can be reconstructed or recovered. Skaloud *et al*. (1999) observed that, accurate priori knowledge of short-term sensor noise distribution is difficult to obtain because of the dependence of their characteristics on other factors such as the technology employed, the manufacturer, environmental factors like vibration, temperature and, tare and ware of the sensors with aging. Therefore, the type of the sensor and the application environment of the navigation sensors will greatly influence the noise level such that the same sensors mounted on aircraft, rail vehicle and road vehicle will have different noise characteristics.

Although several filtering processes are available such as Kalman filtering, Fourier transform and wavelets, their success in GPS and DR integration and error reduction is influenced by accurate prior understanding of the error characteristics. The integrated system models and Kalman filter structures are based on the navigation errors of the component sensors or systems. In this study, the geometric road design elements are recovered purely from the instantaneous curvature values derived from the track position time series. In this section therefore, an investigation is carried out to determine the influences of the deterministic and stochastic errors in the simulated vehicle track positions and curvature computations. The application of cubic smoothing splines in the error reduction model is also examined. This investigation is carried out using simulated data because this presents the opportunity of separating the individual errors.

The analysis of the effects of the errors on the recovery process of the geometric road design elements is done in three parts. The simulated track data free of random and systematic errors is first used in Section (5.2) to test the algorithms developed to estimate the curvature and recover the geometric road design elements. In Section (5.3), the effects of the random noise in distance increments and heading differences on the curvature and position estimation is investigated. The simulated random noise in distance increments were assumed to represent the noise introduced by the wheel speed sensors' measurements as integrated with the measured accelerations. The simulated random noise in heading differences, on the other hand, was assumed to represent the random noise resulting from integration of gyro angular rate measurements. The effects of the simulated systematic errors in distance increments and heading differences are analysed in Section (5.4) while in Section (5.5), the combined effects of both the random noise and systematic errors on curvature and track position estimation are examined. Several systematic errors affecting the inertial sensors and systems are described in

Chapter (3) but only two types are considered in this analysis due to their significance i.e. the distance scale factor error assumed to be analogous to the scale factor error of the ABS wheel speed sensors and the heading bias error assumed to be analogous to bias error in the integrated angular rate gyro measurements. As suggested, the application of cubic smoothing splines to improve on the curvature estimations from the noisy measurements is also analysed.

5.2 The simulated road track

The data simulation in this section is used to test the developed algorithms applied in the extraction of the geometric road design elements and to investigate the effects of the random and systematic errors in the extraction process. The road track coordinates were simulated in a two-dimensional plane as per the geometric road design guidelines specified in *Forschungsgesellschaft für Straßen- und Verkehrswesen* (1995) discussed in Chapter (2). The simulated vehicle route comprises of three types of geometric road design elements: the straight lines, circular curves and transitional curves. The simulated track consist of four straights each of length *1500 m, 1000 m, 400 m* and *400 m* respectively and three circular curves each of radius *585m, 500m* and *550 m* respectively (*cf.* Figure 5.1). The three circular arcs and the four straight lines are joined using six transitional curves each of length either *200 m* or *300 m*. The generated *x* and *y* coordinates form a local right hand coordinate system and the total length of the simulated road track section was about seven kilometres with sample points at *10 m* interval. Figure (5.2) shows the plot of *x* and y coordinates of the simulated track as functions of the track length. The vehicle dynamics are ignored in the simulated data but at an assumed constant speed of *10 m/s*, the total sample time is assumed to be *729 seconds*.

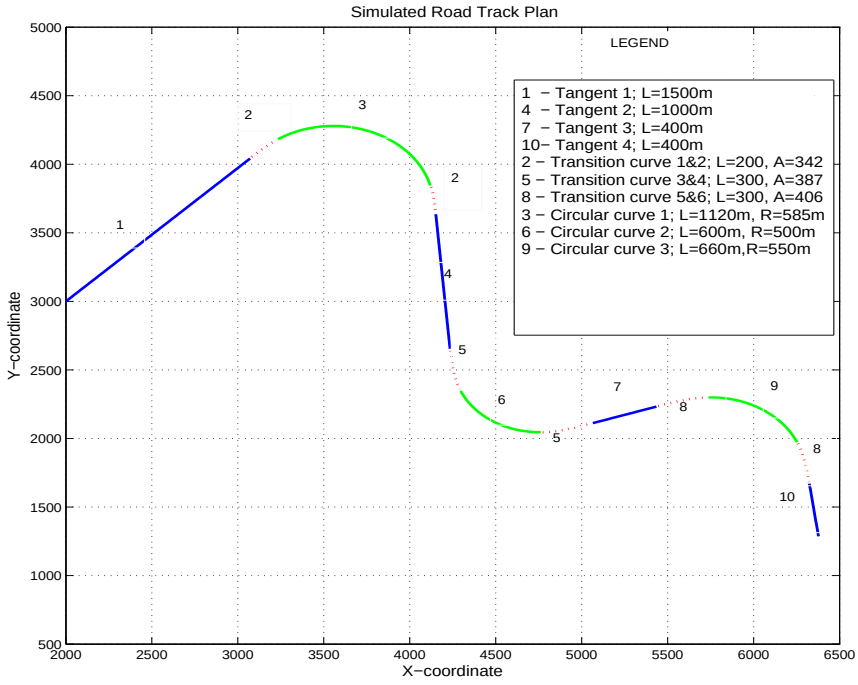

Figure 5.1*: The simulated horizontal geometry vehicle track*

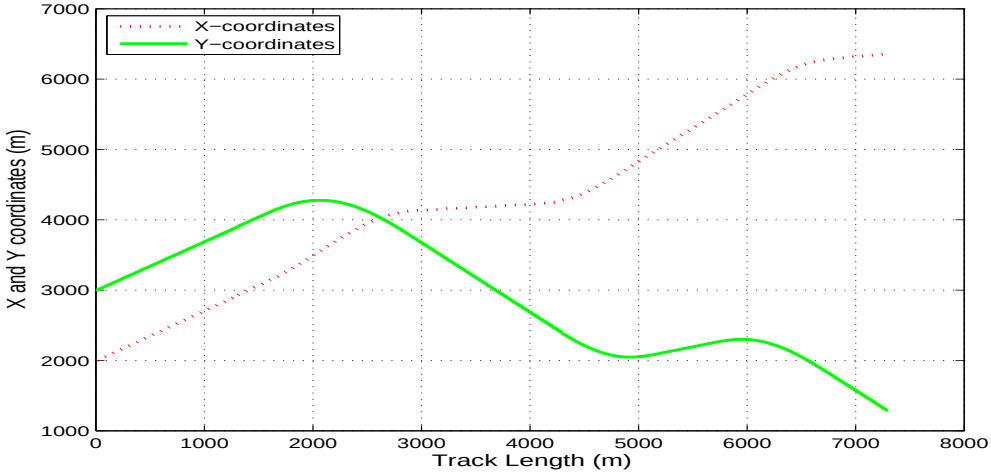

Figure 5.2*: The simulated X and Y increments of the vehicle track*

The simulated track positions, free from random and systematic errors, are used to test the algorithm developed for the recovery of the geometric road design parameters. The algorithm is based on cubic spline interpolation and is described in detail by the flowchart in Appendix (3) and the formulation in section (4.3.3.4). The differences in *x*-coordinates and the first divided differences of the y-coordinates are first computed as shown in Table (4.3). The slopes are then derived and solved using the Gaussian elimination method. By setting the boundary conditions, the spline coefficients are obtained depicting the functional relationship between the data points. Curvatures values along the trajectory are estimated from the first and second derivatives of the spline functions and then analysed to determine the geometric elements i.e. circular curve, transitional curve or straight tangent. The curvature values are further used to determine the start and end of each geometrical road design elements. The comparison of the known geometric design elements shown in Figure (5.3) and the recovered geometric design elements shown in Figure (5.5) gives the error results in testing of the algorithm as shown in Figure (5.4). From the figures, one observes that in the absence of random errors in the measurements and as long as there are no discontinuities in curvature at joining points, the algorithm recovers the geometric road design elements with the accuracies of $\pm 0.5\,m$ in radius and $1.5654\times10^{-6}\ m^{-1}$ in curvature.

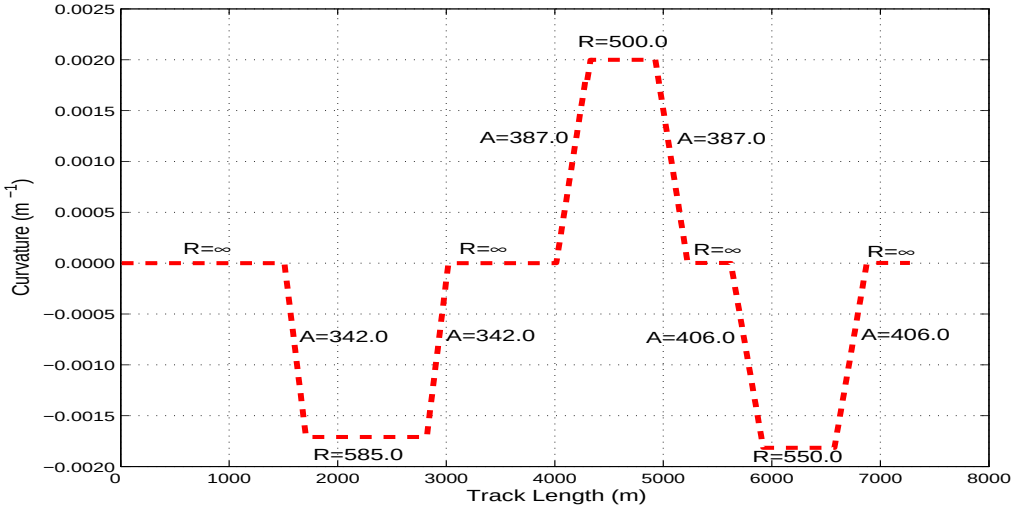

Figure 5.3: *True curvature of the simulated road track.*

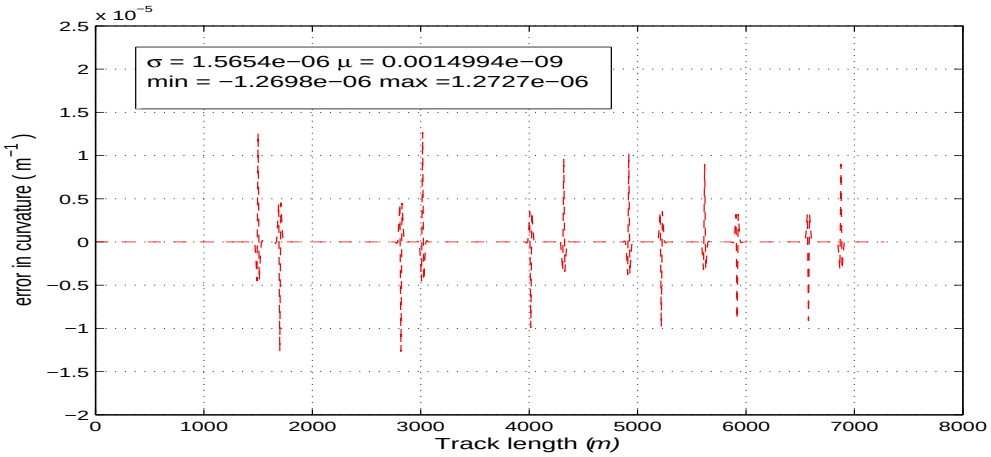

Figure 5.4: *Errors in the vehicle track curvatures*

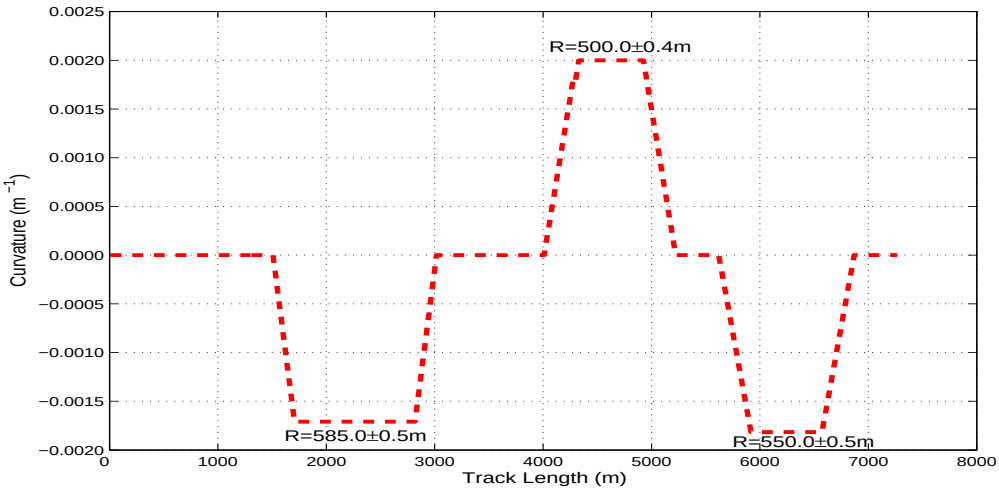

Figure 5.5: *Effects of the errors in curvature on the recovered radii*

5.3 Investigation of the influences of sensor random errors

The vehicle track positions measured using low-cost integrated sensors are subject to significant random and systematic errors. Based on the characterization of the random errors in the field measurements collected using the proposed low-cost sensors as described in Chapter (6), the random and systematic errors are simulated and then introduced into the simulated vehicle track positions. The investigations into the effects of the sensors random noise on the curvature computation using the cubic spline based algorithm and the subsequent recovery of the geometric road track design elements is presented in this section. The analysis and investigations have been carried out separately for the distance increments as detailed in Section (5.3.1) and the vehicle heading differences as contained in Section (5.3.2). As already mentioned in Section (5.1) the simulated random errors in distance increments are assumed to be analogous to the random errors in the distance increments derived from ABS and accelerometer integration while the simulated random errors in headings differences are assumed to be analogous to the random errors in the integrated angular rate gyroscope measurements after error compensations. The combined effects of the random errors in both the distance increments and heading differences are investigated in Section (5.3.3).

5.3.1 Effects of the random noise in distances

The ABS wheel speed sensors always work under severe conditions such as high temperatures, high humidity, vibrations, electromagnetic interference and pollution. The characteristics of the disturbances and interferences to the ABS wheel speed sensors are normally unknown i.e. the signals are buried in a broad band noise background. But there are efficient estimation methods used to determine the magnitude and distribution of the noise levels. For example, robust and optimal filtering is normally used to eliminate noise and interferences in wheel speed sensors. On the other hand, the accelerometer derived distances are known to deviate or drift significantly within a short period, if unchecked, due to accelerometer noise. Although the proposal is to use distance increments derived from the integration of ABS and accelerometer, the assumption in the simulation of the noise at this stage is that the random noise in distance increments are mainly due to the random errors in ABS wheel speed sensors. The investigation therefore determines the influence of the noise on the position error and subsequently on the computation of the curvature values. The fundamental relation of the simulated distance increments and heading differences to the coordinates is given in Equation (5.1) below. Assuming there are no random and systematic errors then;

$$x_{i+1} = x_i + \Delta d_i \sin\left(\theta_i + \Delta\theta_i\right) \quad \text{and} \quad y_{i+1} = y_i + \Delta d_i \cos\left(\theta_i + \Delta\theta_i\right) \tag{5.1}$$

where Δd_i is the simulated distance difference, $\Delta\theta_i$ is the simulated heading difference, θ_i is the simulated vehicle heading and x_i, y_i are the instantaneous vehicle positions at epoch i. The Gaussian white noise in distances is simulated with a mean of zero and standard deviation of $\pm 0.05m$ which is then added to the distance increments Δd_i. The vehicle position contaminated with noise was then estimated as shown in Equation (5.2).

$$x_{i+1} = x_i + \Delta\tilde{d}_i \sin\left(\theta_i + \Delta\theta_i\right) \quad \text{and} \quad y_{i+1} = y_i + \Delta\tilde{d}_i \cos\left(\theta_i + \Delta\theta_i\right) \tag{5.2}$$

where $\Delta\tilde{d}_i = \Delta d_i + \varepsilon_i$ and ε_i are the normally distributed random noise.

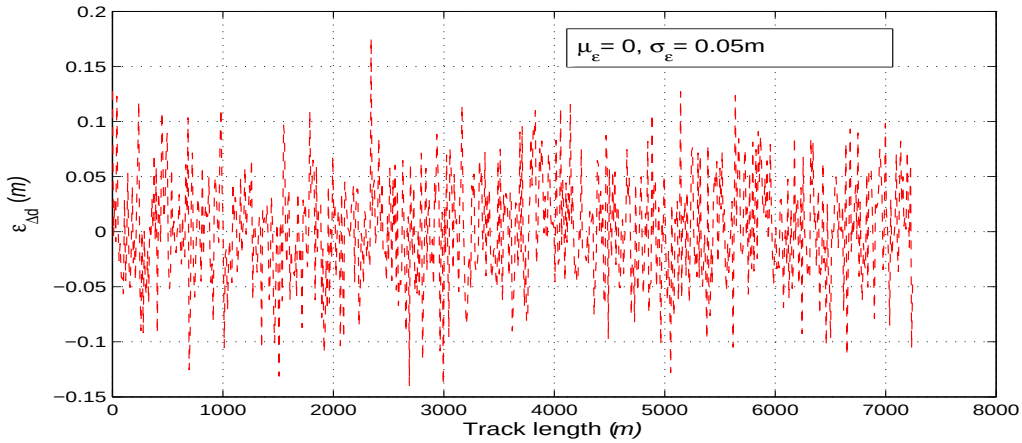

Figure 5.6*: Simulated random noise in distance increments.*

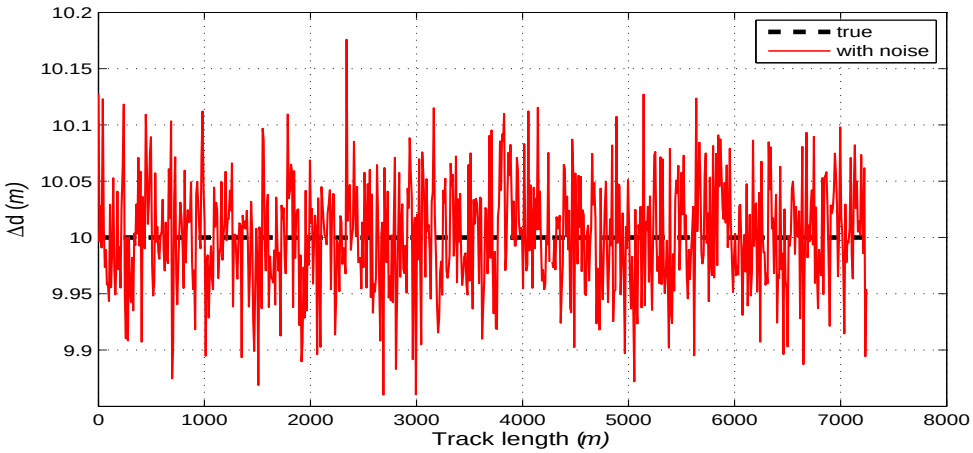

Figure 5.7*: Simulated distance increments (both true and with noise contamination)*

Figure (5.7) shows the true simulated distance increments as well as the distance increments contaminated with random noise. The effects of the random errors in distance increments on the *x*- and *y*-increments of the simulated vehicle track positions and the mean position errors are plotted in Figure (5.8). From the figure, it is observed that the errors in the *x*- and *y*-increments of the simulated vehicle track positions have a standard deviation of about ± 0.035 *m* but remained within the centimetre level and therefore significantly corresponds to the introduced random errors in distance increments. The mean position errors can be concluded to be directly correlated to the random noise in distance increments.

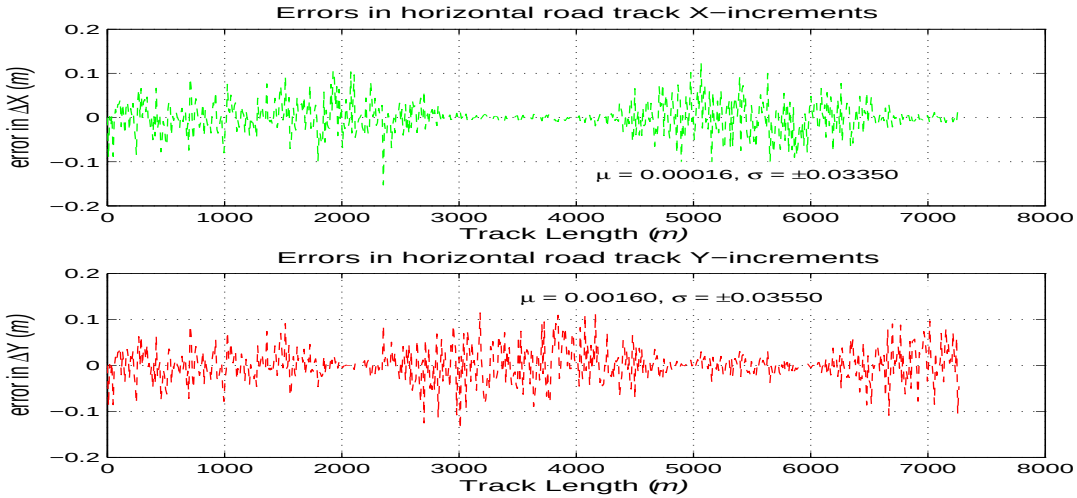

Figure 5.8: *Errors in x and y increments due to random noise in distances*

Obviously the random errors in distance do not affect the curvature estimation along the straight portion of the vehicle track as can be seen in Figure (5.9) showing the errors in curvature due to random noise in distance increments. The errors in curvature due to the random noise in distance increments are though significant in the circular and transitional curve sections of the vehicle track with a mean standard deviation of about $\pm 4.374 \times 10^{-6}$ m^{-1}. The errors in curvature due to the random noise in distance increments translates into the following errors in the recovered radii of curvatures as shown in Figure (5.10); ± 1.5 *m*, ± 1.1 *m* and ± 1.3 *m* in the three circular curve radii of *585 m*, *500 m*, and *550 m* respectively.

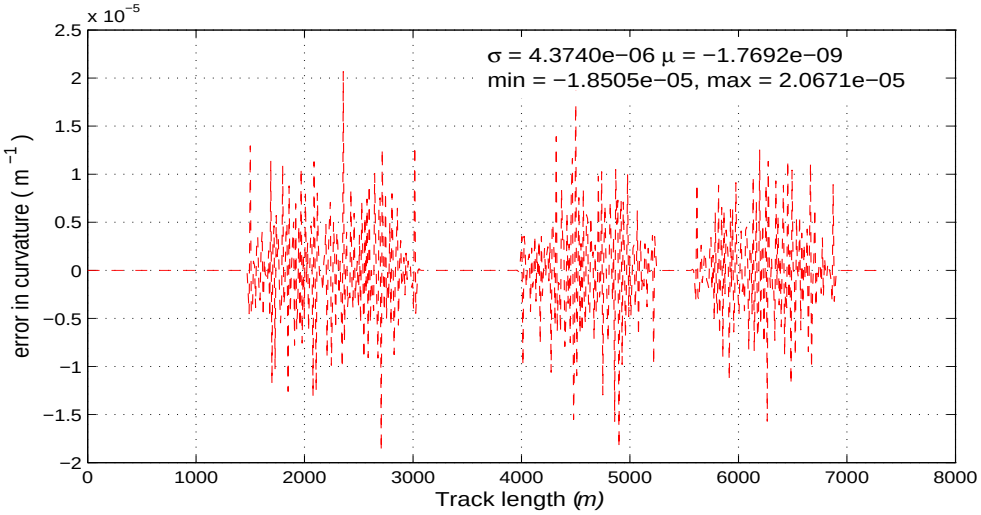

Figure 5.9: *Errors in curvature due to distance random noise*

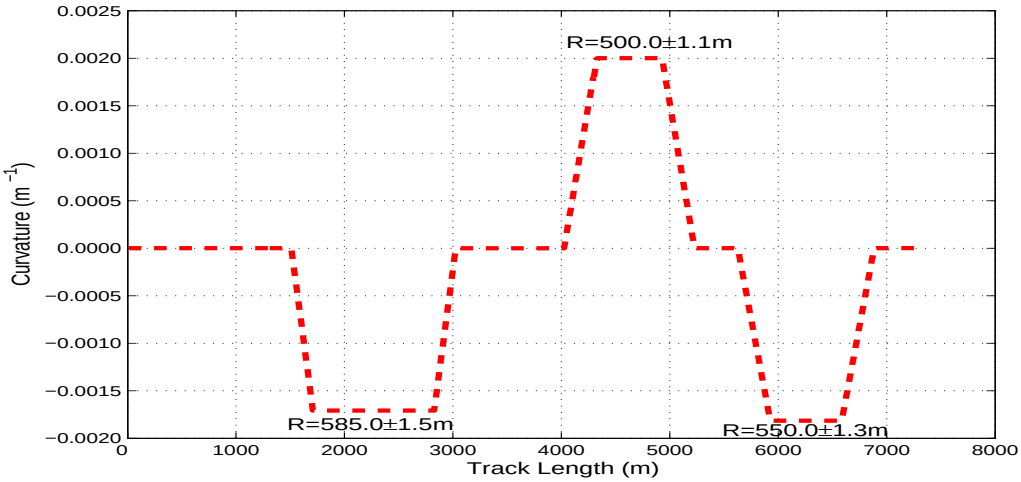

Figure 5.10*: Effects of distance random noise on recovered parameters*

5.3.2 Effects of the random noise in heading

Often the random noise in the rate sensing gyros is specified in degrees per second or as a function of power spectral density (PSD) and even at times as a fast Fourier transforms (FFT). Sometimes the random noise characteristic of the gyroscope is stated as degree per square root hour of a gyro random walk, also referred to as angle random walk (ARW). If one considers the output of a single axis rate gyro with no errors or noise arising from the coupling of the axes and assuming that there is no misalignment error (axis is perpendicular to the gyro spin axis), then the heading angle, θ can be obtained by Equation (5.3) as

$$\theta = \int_{t_0}^{t_0+\Delta t} \omega(t)\, dt \tag{5.3}$$

where $\omega(t)$ is the angular velocity along the sensitive axis. Obviously the errors in heading will exhibit a random walk characteristic over time in the integration process due to the random errors in angular rates whether the sensor is moving or at stand still. Equation (5.1)

can be rewritten as Equation (5.4) assuming only the gyroscope angular rate measurements contributes the random noise.

$$x_{i+1} = x_i + \Delta d_i \sin\left(\theta_i + \Delta\tilde{\theta}_i\right) \quad \text{and} \quad y_{i+1} = y_i + \Delta d_i \cos\left(\theta_i + \Delta\tilde{\theta}_i\right) \tag{5.4}$$

where $\Delta\tilde{\theta}_i = \Delta\theta_i + \gamma_i + \alpha_i$, and α_i are white noise process with $E\{\alpha_i\} = 0$ for all i, i being the epoch. γ_i are the time dependent gyro drift (whose time derivative is purely random noise) and undergoes a *random walk*. On the same principle that Pappas *et al*. (1999) gives a theoretical correlation between the gyro voltage output and the angular velocity output, we can also relate the angular velocity to heading difference as shown in Equations (5.5) to (5.7) below. The gyro voltage output, $g(t)$ is proportional to angular velocity $\omega(t)$ and are related by Equation (5.5) as

$$g(t) = a\omega(t) + b + \varepsilon_g \tag{5.5}$$

where a is the scale (gain) and b is the bias offset. Then

$$\omega(t) = \frac{g(t) - b}{a} + \varepsilon_\omega \tag{5.6}$$

The heading $\tilde{\theta}$ are then obtained by integrating the angular velocities as shown in Equation (5.7).

$$\hat{\theta}(n) = \int_{t_0}^{t_n} \omega(t)\, dt = T\sum_{t_0}^{t_n} \frac{g(t) - b}{a} = \theta(n) + \varepsilon_\theta(n) \tag{5.7}$$

Assuming there is no scale factor error and that the bias offset is zero, the standard deviation of the heading is then estimated using Equation (5.8)

$$\sigma_{\hat{\theta}}(n) = \sigma_\omega T\sqrt{n} \tag{5.8}$$

that is, if ε_ω is random white noise with a standard deviation, σ_ω. T is the sampling period. Sample of the simulated time varying standard deviations in the heading differences as a result of the random noise in angular rates is shown in Table (5.1).

Table 5.1*: Angle random walk (ARW) and gyro Gaussian noise generation*

$ARW = 0.035^o / \sqrt{hr}$			
Sampling Period (s)	Standard Deviation (σ_ω) (°)	Gaussian Noise (ε) (°)	Std. Dev (σ_θ) (°)
1	$0.035/\sqrt{3600} \times \sqrt{1} = \pm 0.000583$	$\varepsilon \sim N(0, 0.0000003)$	0.000184
10	$0.035/\sqrt{3600} \times \sqrt{10} = \pm 0.001845$	$\varepsilon \sim N(0, 0.0000034)$	0.001845
100	$0.035/\sqrt{3600} \times \sqrt{100} = \pm 0.005833$	$\varepsilon \sim N(0, 0.0000340)$	0.018447
729	$0.035/\sqrt{3600} \times \sqrt{800} = \pm 0.015750$	$\varepsilon \sim N(0, 0.0002481)$	0.134476

Figure (5.11) shows the simulated angle random walk obtained by integrating the gyro random noise with respect to curve length and the final deviation in the heading difference is about *0.14°*. This was based on the assumed constant simulated data interval of *1 Hz*. The simulated heading angle random walk is plotted in Figure (5.11). The effects of the random errors in heading differences on the *x*- and *y*-increments of the simulated vehicle track positions are plotted in Figure (5.12). From the figure, the errors in the *x*- and *y*-increments of

the simulated vehicle track positions have a standard deviation of about $\pm 0.005m$ which is within the centimetre level due to the introduced heading random errors of $0.1°$. The mean standard deviation is about $0.007\ m$. with minimum values of $0.008\ m$ and $0\ m$ and maximum values of $0.025\ m$ and $0.025\ m$ in the x- and y-increments respectively.

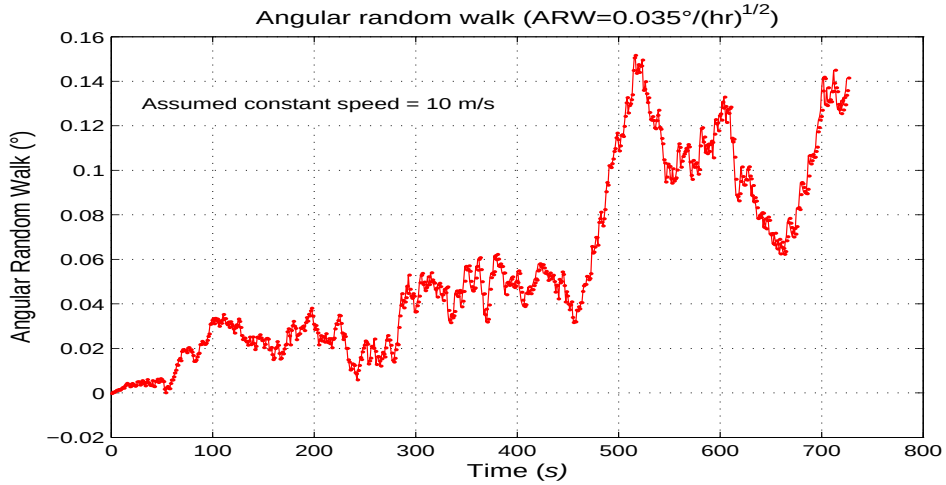

Figure 5.11: *Simulated angular random walk*

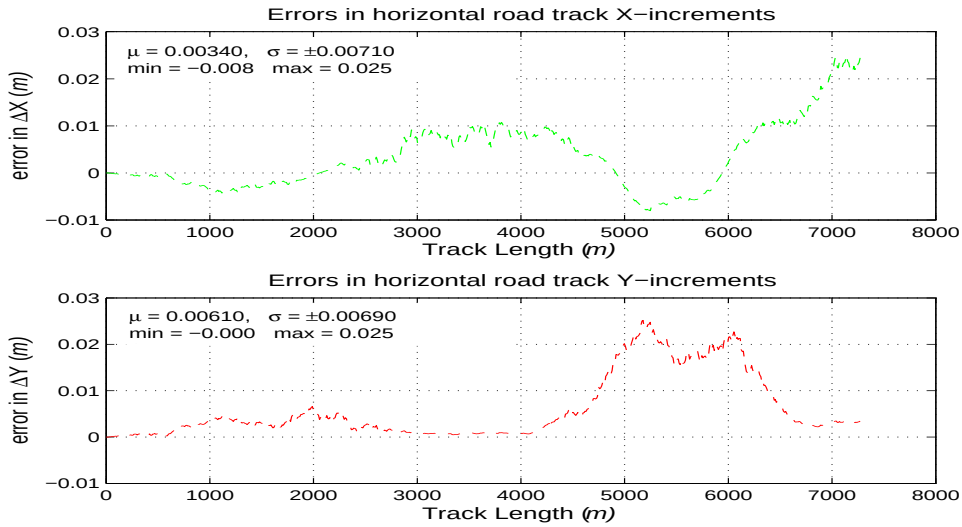

Figure 5.12: *Errors in x and y increments due to heading random noise*

The random errors in heading differences unlike the random errors in distance increments affect the entire vehicle track and increases with time. This is shown in Figure (5.13), which is a plot of the errors in curvature due to random noise in heading differences as functions of the curve length. This is due to the fact that the random noise in heading difference causes both translation and rotation error to the vehicle track. The errors in curvature due to the random noise in heading differences are not significantly different from those due to distance increments with a mean standard deviation of about $\pm 6.126x10^{-6}\ m^{-1}$. The errors in curvature due to the random noise in heading differences translates into the following errors in the recovered radii of curvatures as shown in Figure (5.11); $\pm 2.1m$, $\pm 1.5m$ and $\pm 1.8m$ in the three circular curve radii of $585m$, $500m$, and $550m$ respectively.

As can be observed from the analysis of the random errors in both distance increments and heading differences, the random errors in heading are dominant due to the fact that they cause both translational and rotational position shifts along the simulated vehicle track.

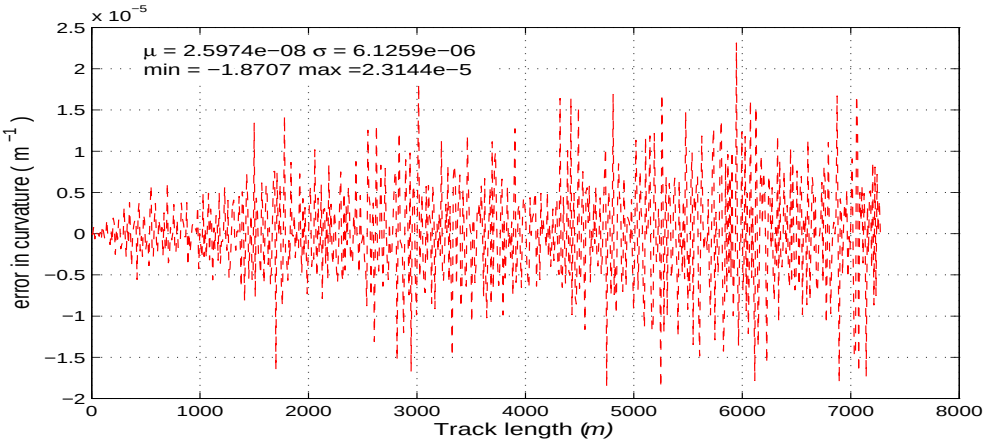

Figure 5.13*: Errors in curvature due to heading random noise*

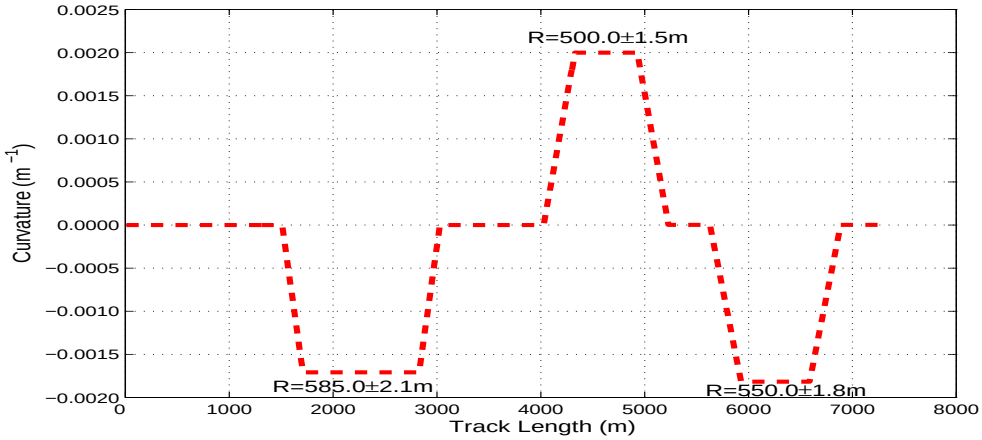

Figure 5.14: *Effects of heading random noise on recovered parameters*

5.3.3 Combined effects of random noise in distance and heading

In this section, the combined effect of random errors in distance increments and heading differences on the simulated vehicle track position and curvature is investigated. This is basically a combination of Equations (5.2) and (5.4) which can be reformulated as

$$x_{i+1} = x_i + \Delta\tilde{d}_i \sin\left(\theta_i + \Delta\tilde{\theta}_i\right) \quad \text{and} \quad y_{i+1} = y_i + \Delta\tilde{d}_i \cos\left(\theta_i + \Delta\tilde{\theta}_i\right) \tag{5.9}$$

where $\Delta\tilde{\theta}_i = \Delta\theta_i + \gamma_i + \alpha_i$, $\Delta\tilde{d}_i = \Delta d_i + \varepsilon_i$ and ε_i, α_i, γ_i are defined as before.

The effects of the combined random errors in distance increments and heading differences on the *x*- and *y*-increments of the simulated vehicle track positions are shown in Figure (5.15). The mean standard deviation is about *0.035 m*. with minimum values of *-0.152 m* and *-0.139 m* and maximum values of *0.120 m* and *0.116 m* in the *x*- and *y*-increments respectively.

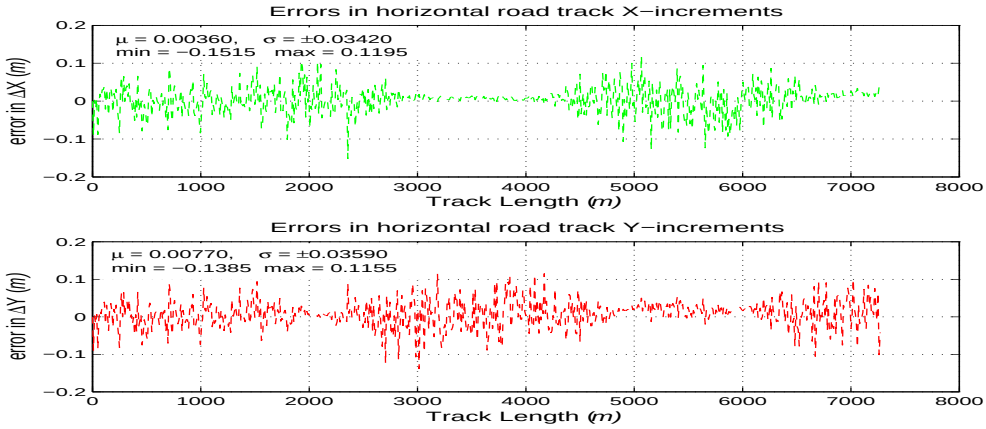

Figure 5.15: *Errors x- and y-increments due to combined distance and heading random noise*

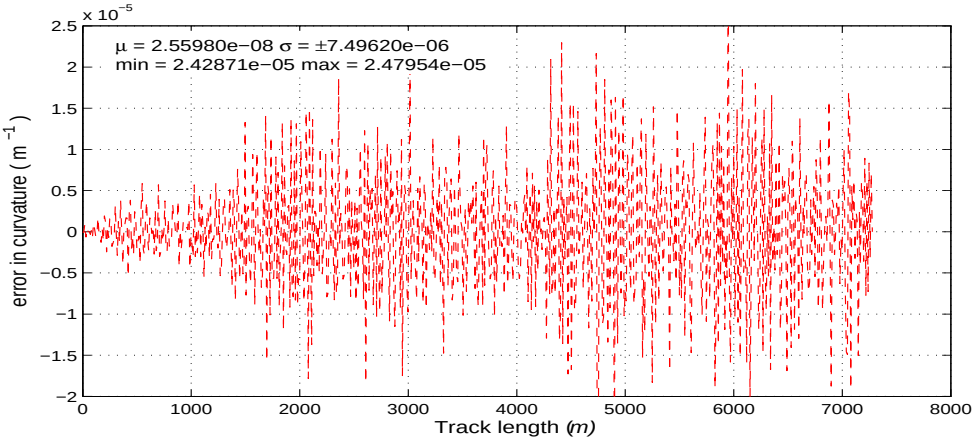

Figure 5.16*: Errors in curvature due to combined heading and distance random noise*

The combined random errors in distance increments and heading differences, similar to the random errors in heading differences, affect the entire vehicle track and increases with time as can be observed in Figure (5.16). The errors in curvature due to the combined random noise are relatively higher in comparison to those due to distance increments or heading differences with a mean standard deviation of about $\pm 7.496x10^{-6}\ m^{-1}$. The errors in curvature due to the combined random noise in distance increments and heading differences translates into the following errors in the recovered radii of curvatures as shown in Figure (5.18); $\pm 2.1m$, $\pm 1.9m$ and $\pm 2.3m$ in the three circular curve radii of *585m*, *500m*, and *550m* respectively.

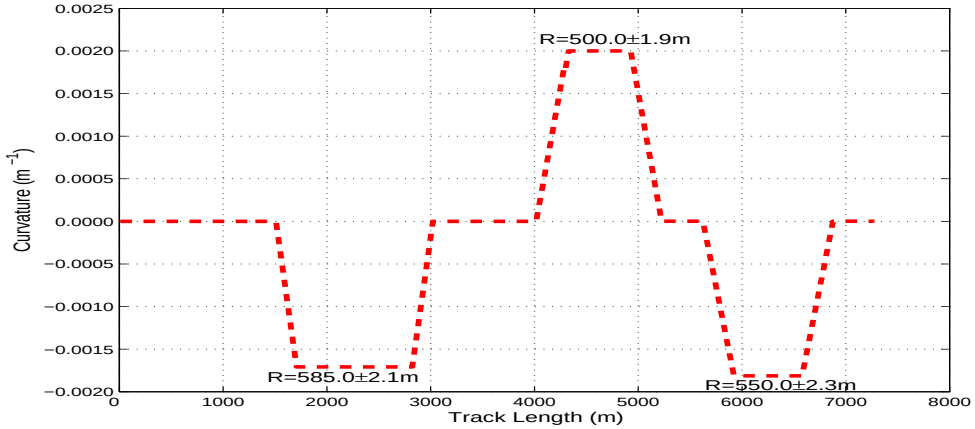

Figure 5.17*: Effects of both distance and heading random noise on recovered parameters*

5.4 Investigation of the influences of systematic errors

The systematic errors are assumed to be analogous to the residual errors exhibited by the installed gyros and ABS wheel speed sensors. Unlike the stochastic errors, they can be characterized and reduced through the calibration process. In this investigation, the assumption is that the errors are not calibrated and their influences on the geometric road design parameters estimation are therefore examined. The relation of the systematic errors to the distance increments and heading differences can also be based on the fundamental expression shown in Equation (5.1).

5.4.1 Effects of scale factor error in distance

The number of signal counts per wheel revolution as measured by the wheel speed sensors can be converted to distance increments travelled by the vehicle through multiplication by a conversion factor, hereafter referred to as scale factor, obtained from an initial calibration process. The scale factor of the wheel speed sensor can be optimally estimated by driving the vehicle over a known distance and applying a Kalman filter algorithm as described in Chapter (6). The scale factor is never constant within the measuring session and changes over time and from wheel to wheel mainly due to the wheel slippage and skidding, the tire pressure variation, the tire wear and the vehicle speed variation along track during the simulation. The variation of the scale factor error is assumed to be insignificant. The error is therefore considered to be a constant with a bias of *0.2%* of the full scale factor.

Using similar notations as in Equation (3.6) where the scale factor error e is expressed in percentage of the full true scale factor s, then the simulated scale factor $\tilde{s}$ contaminated by scale error is given as by Equation (5.10)

$$\tilde{s} = s + \frac{es}{100} \tag{5.10}$$

Then, the distance increment contaminated with scale factor error $\Delta\tilde{d}$ is given by

$$\Delta\tilde{d} = \tilde{s}.\Delta p = \left(s + \frac{es}{100} \right) \Delta p = \Delta d + \frac{e}{100} \Delta d \tag{5.11}$$

Equation (5.11) shows that the scale factor error in distance increment depends on the magnitude of the distance increment travelled by the wheels. The error therefore depends on the speed of the measuring vehicle, if the sampling interval is constant. The simulated distance increments are assumed to be constant i.e. constant speed. The distance increments travelled by the wheels and contaminated by scale factor error are simulated using Equation (5.11). The scale factor error contributes significant position error as shown in Figure (5.18). The figure shows the cumulated errors in distance due to the sensors scale factor error with respect to curve length resulting reaching a maximum deviation in the distance of about *14.5 m*. The effects of the scale factor error in distance on the *x*- and *y*-increments of the simulated vehicle track positions are shown in Figure (5.19). The mean standard deviation is about *0.455 m*. with minimum values of *-1.330 m* and *-0.375 m* and maximum values of *0.001 m* and *1.488 m* in the *x*- and *y*-increments respectively.

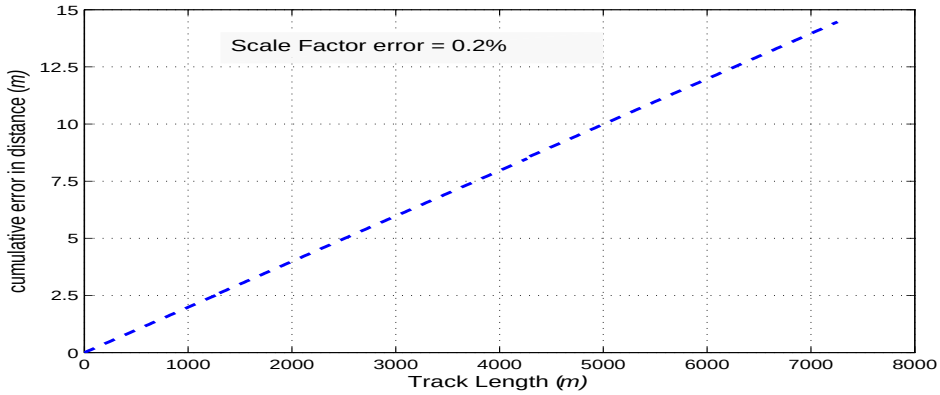

Figure 5.18*: Simulated cumulative error in distance due to scale factor errors*

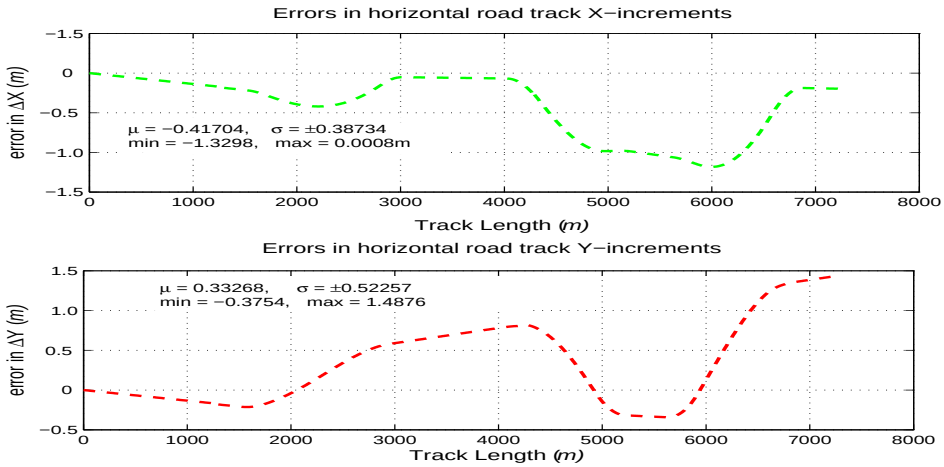

Figure 5.19: *Errors in x and y coordinates due to distance scale factor error*

The scale factor error in distance, similar to the random errors in distance, do not affect curvature along the straight sections of the vehicle track i.e. affects only the transitional and circular curve sections. The error effects on curvature also increases with time as shown in Figures (5.20). The errors in curvature due to the scale factor error in distance increments have a mean standard deviation of about $\pm 2.9271x10^{-6}\ m^{-1}$. The errors in curvature due to the random noise in heading differences translates into the following errors in the recovered radii of curvatures as shown in Figure (5.21); $\pm 0.9m$, $\pm 0.7m$ and $\pm 0.8m$ in the three circular curve radii of *585m*, *500m*, and *550m* respectively.

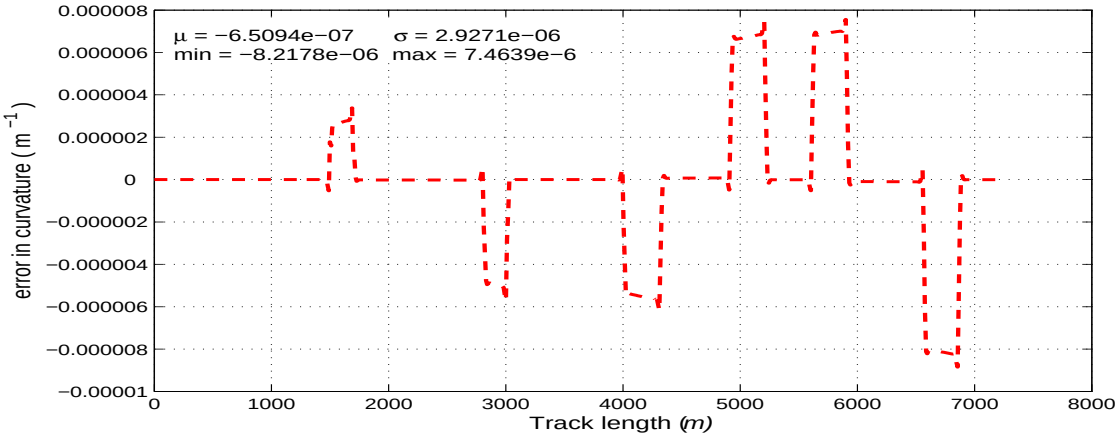

Figure 5.20*: Errors in curvature due to distance scale factor error*

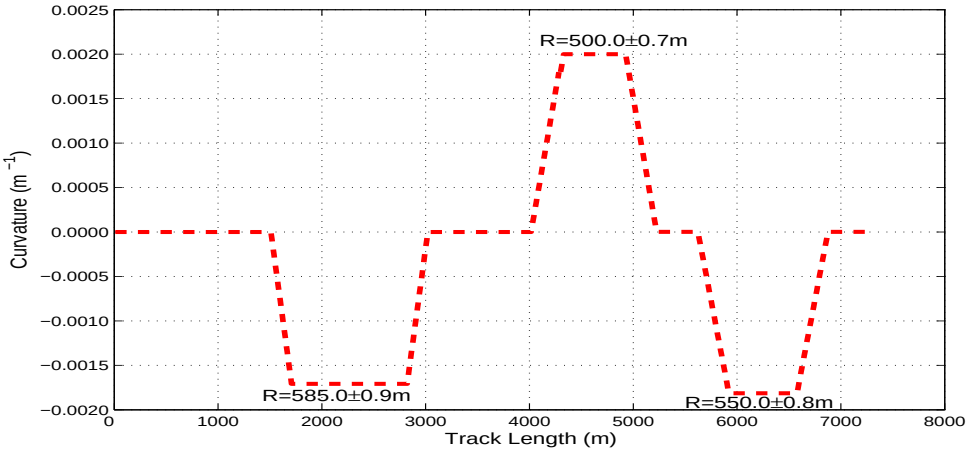

Figure 5.21*: Effects of distance scale factor error on recovered parameters*

5.4.2 Effects of systematic errors in heading

The systematic errors affecting the gyroscope are discussed in Section (3.1.2) and include the gyro bias error and gyro scale factor error. Since the scale factor error in heading differences affects the measurements only when the vehicle is turning, varies with time and changes insignificantly in short intervals i.e. the error has low covariance, only the effects of bias errors in heading differences are investigated in this section. Figure (5.22) shows the effects of the bias error on the simulated heading differences with respect to curve length resulting in a constant deviation in the heading difference of *0.01°*. The effects on the *x*- and *y*- increments are shown in Figure (5.23). The mean standard deviation is about *0.393 m*. with minimum values of *-0.184 m* and *-0.565 m* and maximum values of *1.170 m* and *0.933 m* in the *x*- and *y*- increments respectively. Similar to distance scale factor error, the heading bias effects are significant along the transitional curve sections.

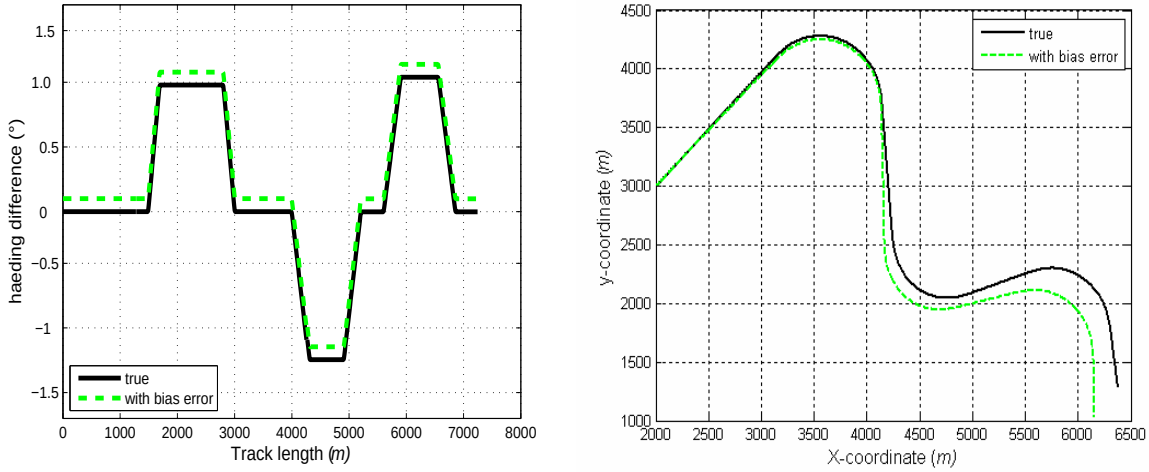

Figure 5.22*: Simulated true and biased heading differences*

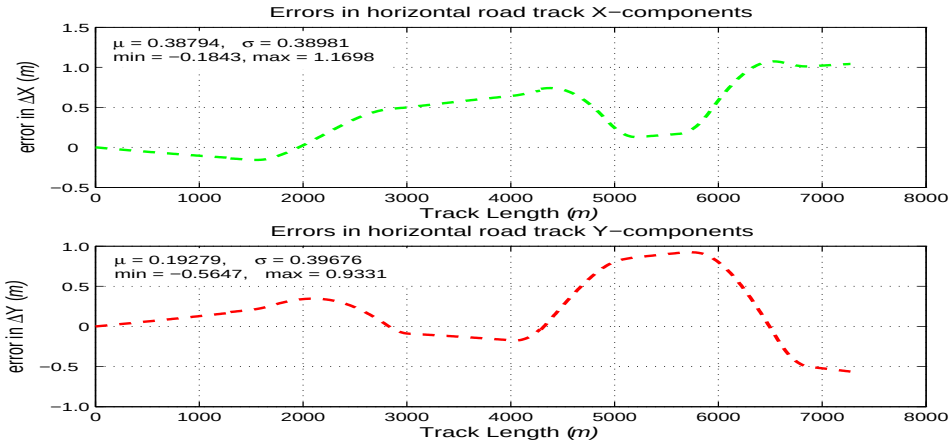

Figure 5.23: *Errors in x and y increments due to heading bias*

The bias error in heading differences like the scale factor error in distance increments do not affects the entire simulated vehicle track. The error causes discontinuities at the connection points as can be seen in Figures (5.24). This is due to the fact that the random noise in headings causes both translation and rotation error to the vehicle track. The errors in curvature due to the bias error in heading differences are therefore not significantly different from those due to scale factor error in distance increments and have a mean standard deviation of about $\pm 1.5934 \times 10^{-6}\ m^{-1}$. The errors in curvature due to the random noise in heading differences translates into the following errors in the recovered radii of curvatures as shown in Figure (5.25); $\pm 1.5\ m$, $\pm 1.1\ m$ and $\pm 1.3\ m$ in the three circular curve radii of *585 m*, *500 m*, and *550 m* respectively. Although the systematic errors can be reduced through calibration process, investigation of their effects is based on the assumption that no calibration process is perfect hence the calibration errors.

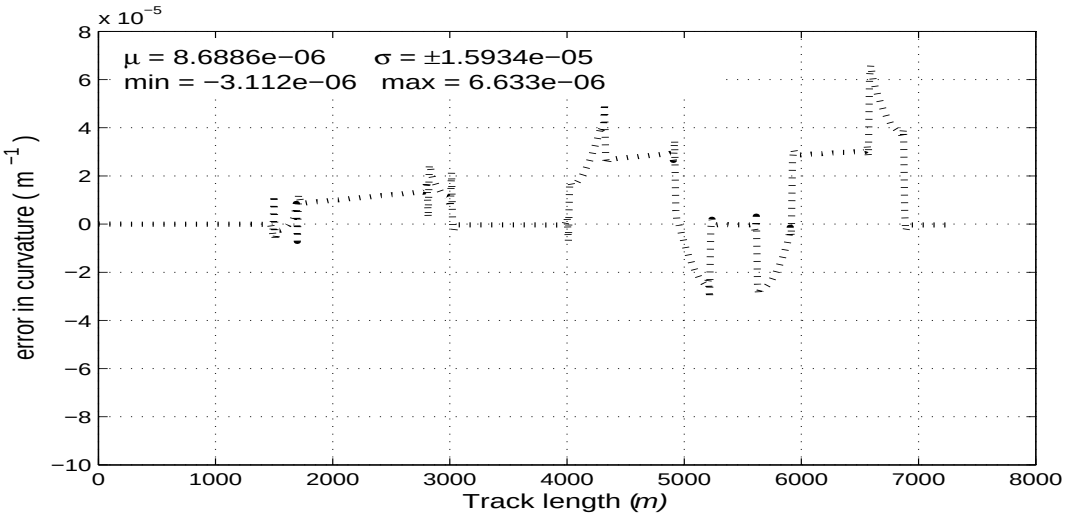

Figure 5.24*: Errors in curvature due to heading bias (0.1°/s)*

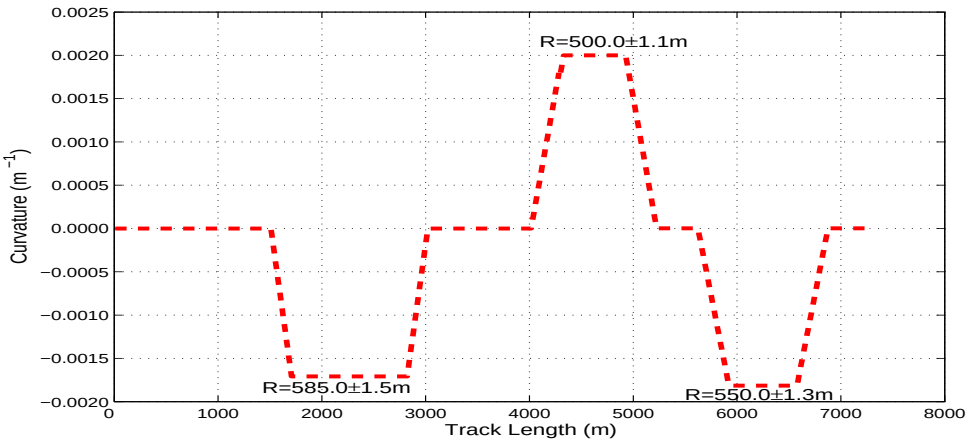

Figure 5.25*: Effects of heading bias error on recovered parameters*

5.5 Influences of the combined random and systematic errors

Investigations into the effects of the combined random and systematic errors on the *x*- and *y*-components of simulated vehicle track positions, mean position and curvature estimation are also carried out. The effect of the combined random and systematic errors on *x*- and *y*-coordinates is shown in Figure (5.27). From the figure, the mean errors in both the *x*- and *y*-increments of the simulated vehicle track positions have a standard deviation of about $\pm 0.7\ m$. The combined effect of random and systematic errors in distance increments and heading differences, like in the case of the combined random errors in distance and heading, affect the entire vehicle track and increases with time as can be seen in Figures (5.28) shows the errors in curvature due to combined random noise in distances and heading differences. Similar to the previous case the random noise in distances and headings causes both translation and rotation errors to the vehicle track.

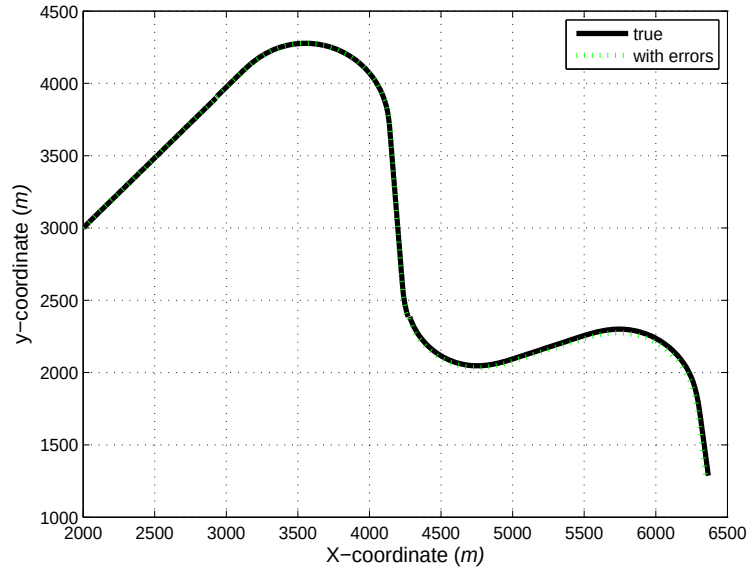

Figure 5.26*: Effects of combined random and systematic errors on track plot*

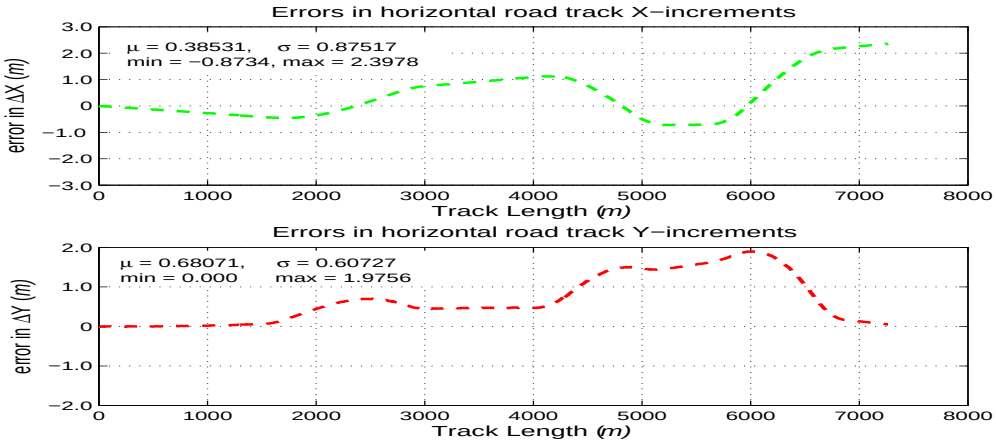

Figure 5.27: *Errors in x and y coordinates due to combined random and systematic errors*

The errors in curvature due to the combined random noise are therefore slightly larger relative to those due to distance increments or heading differences with a mean standard deviation of about $\pm 6.9095x10^{-6}\ m^{-1}$. The errors in curvature due to the combined random noise in distance increments and heading differences translates into the following errors in the recovered radii of curvatures as shown in Figure (5.29); $\pm 2.4\ m$, $\pm 1.7\ m$ and $\pm 2.\ 1m$ in the three circular curve radii of *585 m*, *500 m*, and *550 m* respectively. The conclusion is that the effects of the combined random and systematic errors in distance and heading play a slightly more significant role than the random or systematic errors alone.

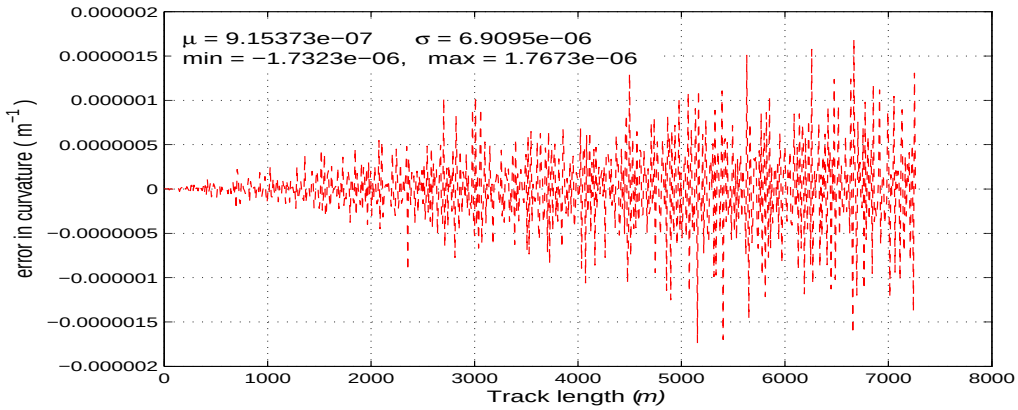

Figure 5.28: *Errors in curvature due to combined random and systematic errors*

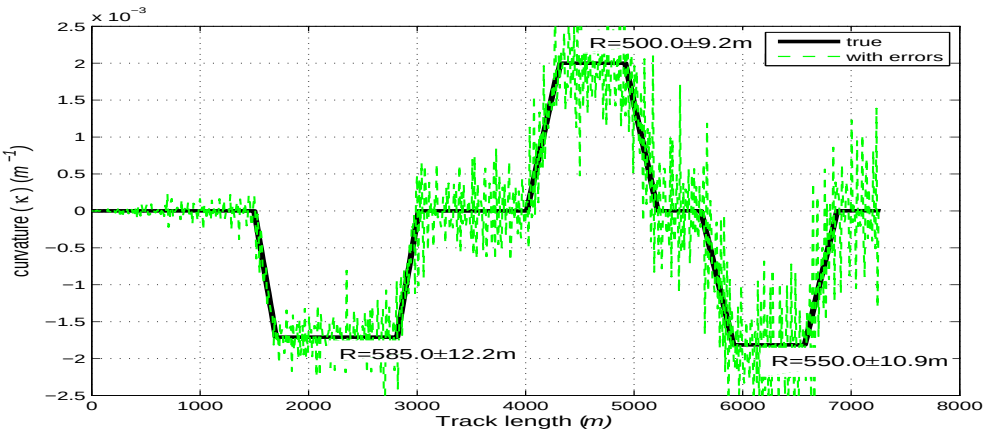

Figure 5.29: *Effects of combined random and systematic errors on curvature*

5.5.1 Effects of data smoothing on combined random and systematic errors

In this section cubic smoothing splines are applied to reduce the effects of random and systematic errors on the road track *x*- and *y*-increments and the estimated curvatures. An analysis of positional errors as shown in Figure (5.30) shows no significant improvement in the errors due to combined random and systematic errors after smoothing. But as shown in Figure (5.31), the curvature errors as a result of the combined effects are reduced drastically.

The effects of the combined random errors in distance increments and heading differences on the *x*- and *y*-increments of the simulated vehicle track positions and the mean position errors are plotted in Figures (5.32). The mean errors in both the *x*- and *y*-increments of the simulated vehicle track positions have a standard deviation of about ± 0.7 *m*.

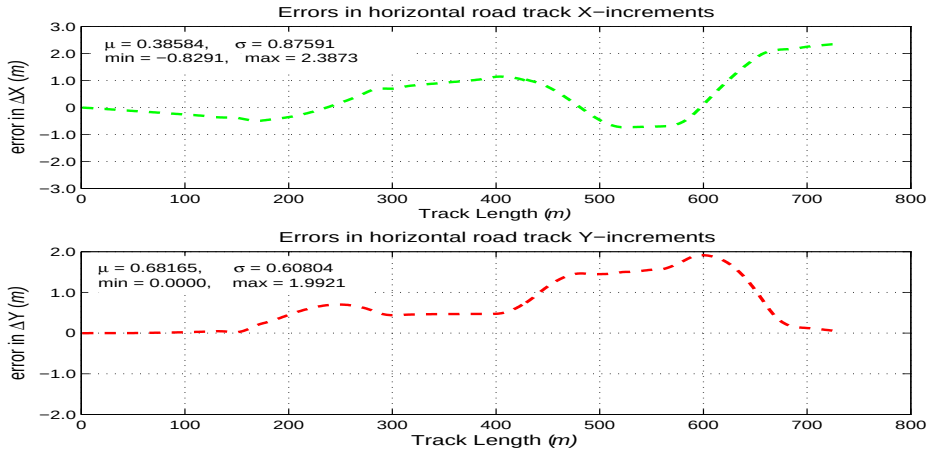

Figure 5.30: *Errors in x and y coordinates due to combined errors after smoothing*

After smoothing, the errors in curvature due to the combined random and systematic errors have a mean standard deviation of about $\pm 4.0276 \times 10^{-6}$ m^{-1}. The errors (after smoothing) in curvature due to the combined random noise in distance increments and heading differences translates into the following errors in the recovered radii of curvatures as shown in Figure (5.31); ± 1.4 *m*, ± 1.0 *m* and ± 1.2 *m* in the three circular curve radii of *585 m*, *500 m*, and *550 m* respectively. In comparison to the results on the investigation of the effects of random noise in Section (5.3) one may conclude that the effects of the random errors in heading differences play a major role than the random errors in distance increments.

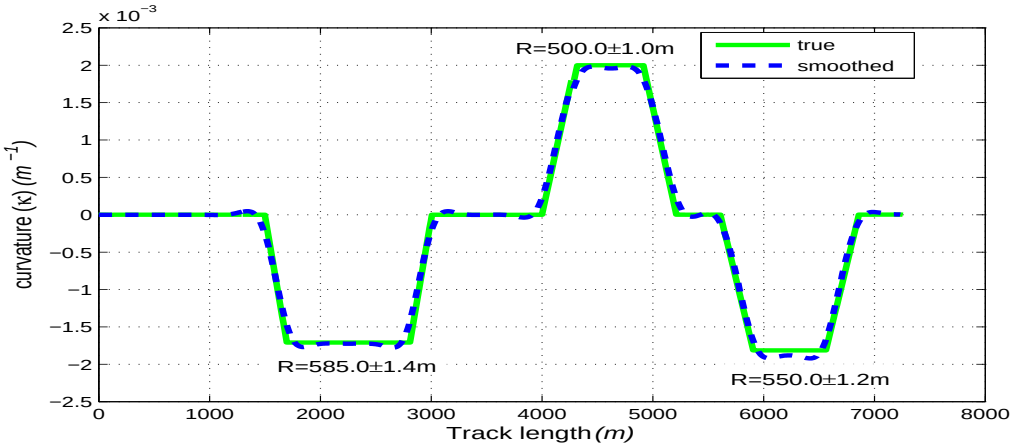

Figure 5.31: *Effects of combined random and systematic errors on curvature after smoothing*

From Figures (5.30) and (5.31), the smoothing spline improves the curvature estimation significantly. An improvement in curvature computation but not on the position i.e. even though the track will be displaced due to combined random and systematic errors, the shape is recoverable using the cubic smoothing curves.

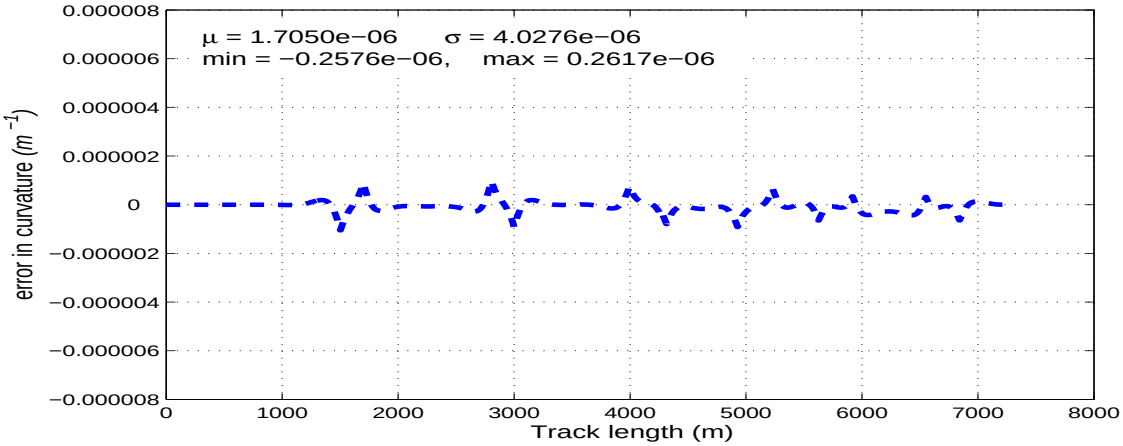

Figure 5.32: *Errors in curvature after smoothing the combined random and systematic errors*

The standard deviations of the estimated position time series and curvatures are summarised in Table (5.2). Analysis of the effects of random errors on the positions show that the random errors in distance affects the positions relatively more than the random errors in heading as shown in Figure (5.33). On the other hand, the smoothing splines do not cause significant change in the accuracy of the estimated *x*- and *y*- increments. As shown in Figure (5.34), the cubic smoothing splines reduce the random error effects, but since the systematic errors predominate, the effect of smoothing in the combined case is negligible.

Table 5.2: *Summary of standard deviations due to random and systematic errors*

	X-increment	Y-increment	Curvature
distance random noise	0.0335	0.0355	0.000004374
heading random noise	0.0071	0.0069	0.000006126
combined random noise	0.0342	0.0359	0.000007496
scale factor error	0.3873	0.5226	0.000002927
bias error	0.3898	0.3968	0.000001593
combined effect	0.8752	0.6073	0.000006909
combined effect (smoothed)	0.8759	0.6080	0.000004028

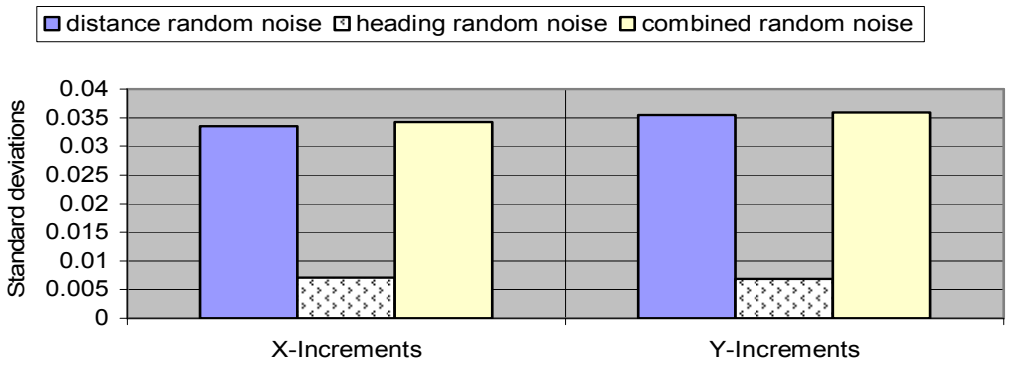

Figure 5.33: *Summary of the effects of random errors on positions*

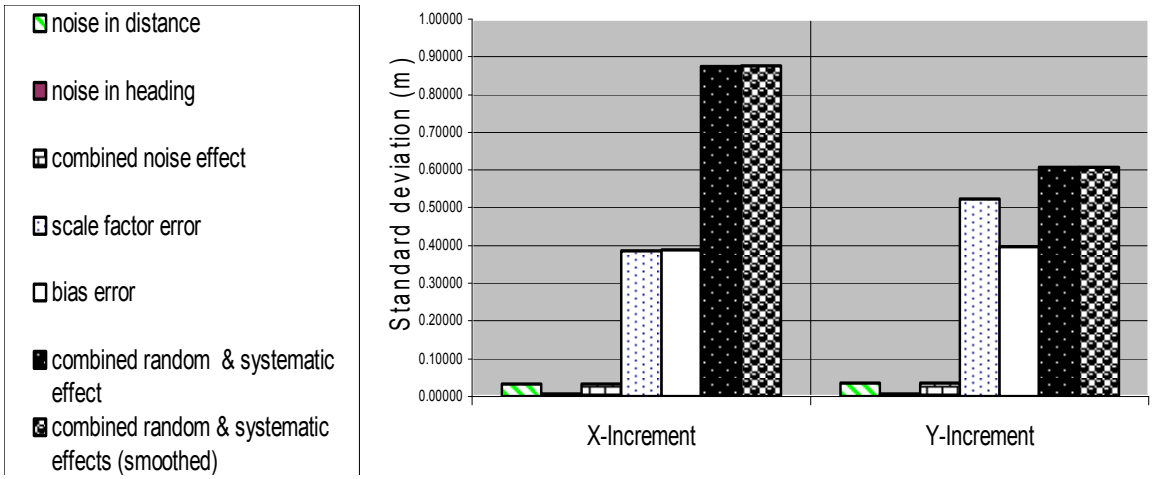

Figure 5.34: *Summary of the effects of systematic errors on positions*

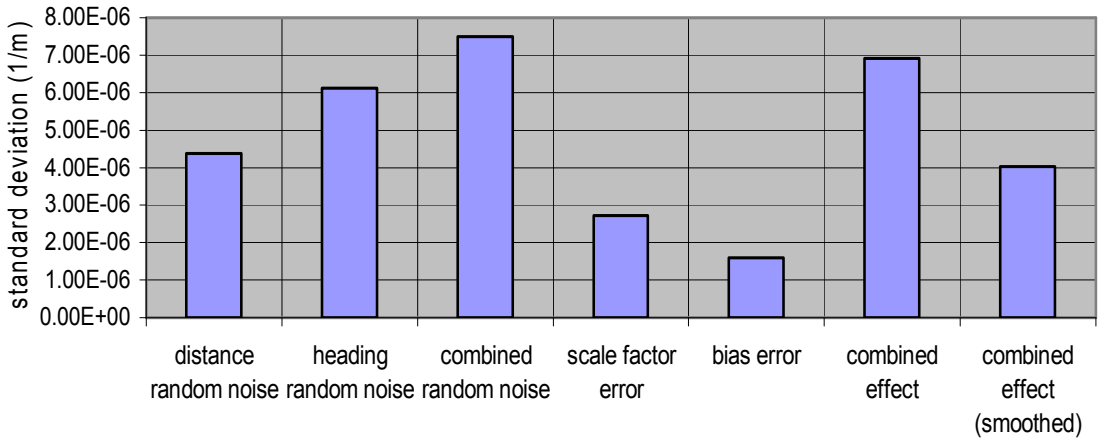

Figure 5.35: *Summary of the effects of random and systematic errors on curvature estimation*

Table (5.3) shows a summary of the range of the errors in curvature as a result of the random and systematic errors in distance increments and heading differences. The range is defined as the difference between the minimum and the maximum errors in curvature. Figure (5.36) shows the analysis of each of the error categories. In comparison to the analysis of the standard errors of the estimated curvature in Figure (5.35), similarity in magnitude can be drawn. The errors in curvature due to the random noise in the distance increments and heading differences are relatively high than those due systematic errors. Though not significantly different, the effects of random noise in distance increments and heading differences is greater than the effects of systematic errors. Similar to the standard deviation analysis, the application of smoothing splines on the combined random and systematic errors has minimal effects on the errors in curvature.

Table 5.3: *Summary of standard deviations due to random and systematic errors*

	Standard deviation	
Type of error	Minimum Maximum	Range (max-min)
Non	-0.000001270 +0.000001273	0.000002543
Distance noise	-0.000016505 +0.000020671	0.000037176
Heading noise	-0.000018707 +0.000023144	0.000041851
Combined noise	-0.000024287 +0.000024795	0.000049082
Scale factor	-0.000008218 +0.000007464	0.000015682
Heading bias	-0.000003112 +0.000006633	0.000009745
Combined	-0.000002429 +0.000002480	0.000004909
Smoothed	-0.000001877 +0.000001362	0.000003239

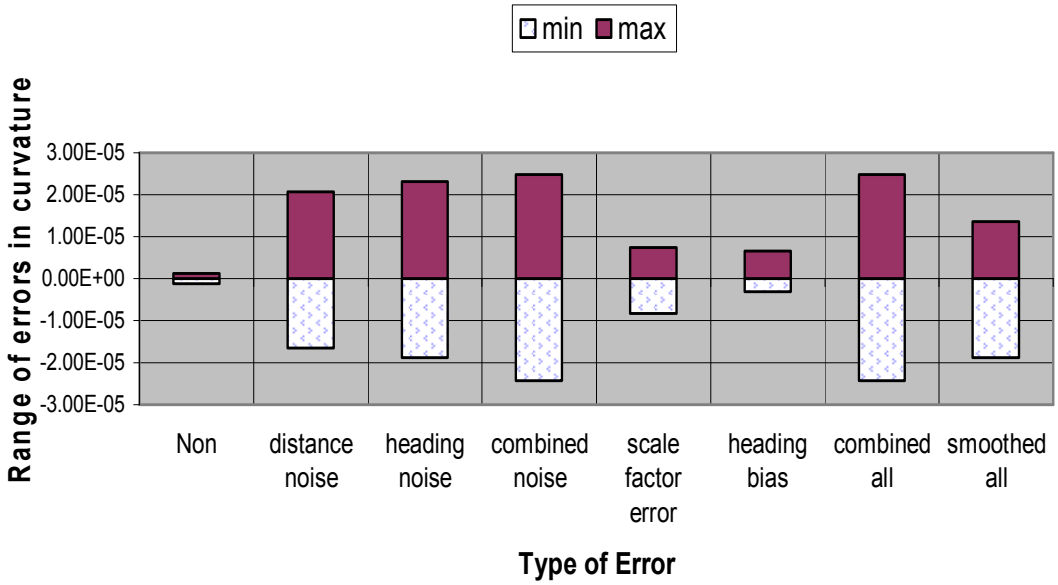

Figure 5.36: *Range of the maximum and minimum standard deviations of the curvature*

6 Tests and Results

Equation Section 6

The applications of navigation sensors in automotive industry require low-cost, small-sized, low-powered inertial sensors. Due to current advances in MEMS technology that enables the development of chip-based low-cost inertial sensors, this requirement has so far been fulfilled. This chapter presents the algorithms and test experiments performed in the implementation of low-cost navigation sensors to recover the geometric road design parameters. The author-developed algorithms run in MATLAB programming language environment. Section (6.1) introduces the proposed low-cost sensors, the field measurements performed and the tests carried out to achieve four main goals. First, Section (6.2) describes the tests and analysis performed for the purpose of synchronising both the ABS wheel speed sensors with the Applanix POS-AV™ derived geometry on one hand and the MotionPak™ measurements with the Applanix POS-AV™ measurements on the other hand. The Applanix POS-AV™ measurements were adopted as the "true values". The formulation and calibration process used to obtain the conversion factors (herein referred to as scale factors) for converting the ABS measured pulse-differences to distance increments and the estimation of threshold speed are also contained in this section. Secondly, Section (6.3) contains the formulations and computations aimed at integrating the ABS derived cumulative distances and the MotionPak™ single-axis acceleration measurements through a Kalman filter process. Thirdly, Section (6.4) presents the processes involved in estimating the vehicle heading from the MotionPak™ single axis gyroscope measurements. Lastly, Section (6.5) details the road track curvature estimation and analysis. In each of the cases where the user developed algorithms were applied, simulated data was first employed in the test and verification of the MATLAB written codes, before they are applied in the processing of the field measured data.

6.1 The low-cost navigation sensors

Measurement of the vehicle track position series and the recovery of the road geometry design elements require an assembly of georeferencing sensors mounted on a vehicle platform to make the necessary measurements of the vehicle dynamics and/or position. A se of appropriate algorithms are also required to analyse the data, eliminate the effects of errors and compute the position time series. The curvature values and the sizes of the design elements are then derived from the position time series.

Figure (6.1) illustrates the equipment set-up for the field research experiments. The low cost sensors consisted of four ABS wheel speed sensors, a single-axis MEMS gyroscope and a single-axis MEMS accelerometer. Note that the accelerometer and the gyroscope are the components of the MotionPak™ system. The wheel speed sensors are connected to a desktop computer running the TURBO PASCAL program for real time data logging. The computer also receives and stores time tag and pulse per second (PPS) signals from the GPS receiver and the wheel speed sensor measurements at *10 Hz*. The MotionPak™ is controlled by a laptop computer and sends event markers to the GPS receiver and the Applanix system at *100 Hz*. The Trimble GPS receiver is used as a back-up system, to generate time take for the ABS sensors and to store event markers for the MotionPak™. The Applanix system comprising of the receiver and the IMU are controlled by a laptop.

The measurements and data processing results obtained from the relatively superior Applanix POS AV™ system are used to compare and analyse the results obtained by the low-cost

sensors. The Applanix POS AV™ system is described in Chapter (3) while the MotionPak™ system is described briefly in Section (6.1.1) below.

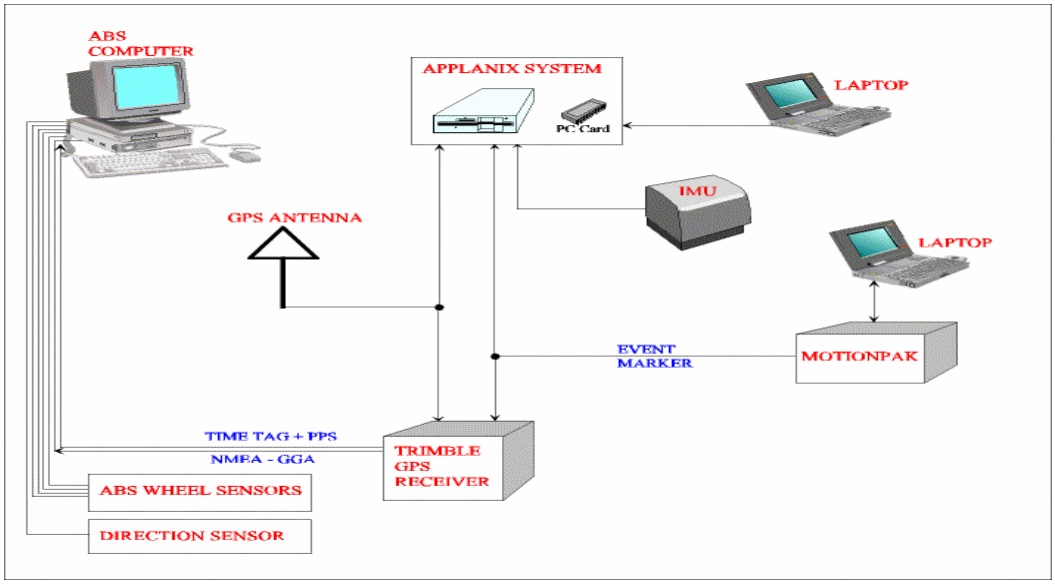

Figure 6.1*: Experimental equipment set-up of the low cost integrated navigation system*

6.1.1 MotionPak™

The MotionPak™ shown in Figure (6.2) is a low-cost solid state six degrees of freedom inertial measuring unit developed by Systron Donner (2006) inertial division of the BEI Technologies Inc. and consists of three micromachined quartz angular rate sensors and three linear servo accelerometers fitted at right angle to one another in a strap-down-system. The axes define the right hand coordinate system. The IMU has an incorporated temperature sensor. The sensors are bound in a *77.5 by 77.5 by 91.5 mm* case with internal power regulation and signal conditioning electronics and weigh about *900g*. The technical parameters of the system provided by the manufacture are as listed in Table (6.1). The MotionPak is design for automotive, aerospace, marine and defence applications and senses linear acceleration and angular rates. The MotionPak™ gyroscope has a bias drift rate of *0.02°/s=1.2°/min=72°/hr*.

Figure 6.2 : *MotionPak™ by* Systron Donner (2006)

The outputs of the inertial measuring unit are sensitive to voltage hence a regulated ± 15 *V DC* power is supplied. The MotionPak includes three micromachined angular rate sensors of type Gyrochip^{TM} model *QRS11* (Quartz Rate Sensor 11), a special gyro for military and aircraft. The rate gyroscopes are capable of measuring angular velocities up to $\pm 200°/s$ while the accelerometers are capable of measuring acceleration up to $\pm 5g$. The three micromachined accelerometers are of type QFA (Quartz Flexible Accelerometer) with a standard range of $\pm 5g$. The temperature sensor is of type Analogue Devices AD 590 measuring temperatures in the range of *-55°C to +150°C*.

Table 6.1: *MotionPak^{TM} performance specifications* (source: Systron Donner, 2006)

Parameter	Rate Channels	Acceleration Channels
Standard ranges	$\pm 200^{o}$C/sec	5g
Input Voltage	± 15Vdc $\pm 10\%$	
Bias variation over Temperature (max. deviation from 22^{o}C	$\leq 3\ ^{o}$/sec from 22 ^{o}C	$\leq 100\ \mu$ g
Bias	$\leq 2\ ^{o}$/sec	$\leq 1000\ \mu$ g
Output Noise (DC to 100Hz)	$\leq\ 0.01\ ^{o}/\text{sec}/\sqrt{Hz}$	≤ 7.0mV

Development of an assembly of low-cost navigation sensors would involve the design of the architecture of the sensors, the design of the signal circuitry and the interfaces, the mounting and alignment in the vehicle and A/D converter not to mention development of the algorithms required or the system calibration and data processing. Since our aim is to use a single axis gyro and a single axis accelerometer, the design, development and implementation of an entirely new navigation sensor is beyond the scope of this work. The commercial-off-the-self IMU (MotionPak^{TM}) and ABS wheel speed sensors was therefore used to measure the position and direction change information. Although the MotionPak is six degrees of freedom system, only one channel of the gyro (*z*-axis angular rate) and one channel accelerometer (*x*-axis acceleration) is adopted. More information on the MotionPak^{TM} system is contained in Section (6.1.2) and Systron Donner (2006).

6.1.2 Anti-lock braking systems (ABS)

The anti-lock braking system, commonly known as the ABS system or simply as ABS, is described in section (3.2) and consists of three main components, of interest being the wheel speed sensors. The wheel speed sensors used in the proposed low-cost sensors set-up are fitted in all the four wheels of the measuring van (*Volkswagen LT 35*) and uses passive magnetic variable reluctance technique described in Section (3.2) to sense the wheel rotation speed.

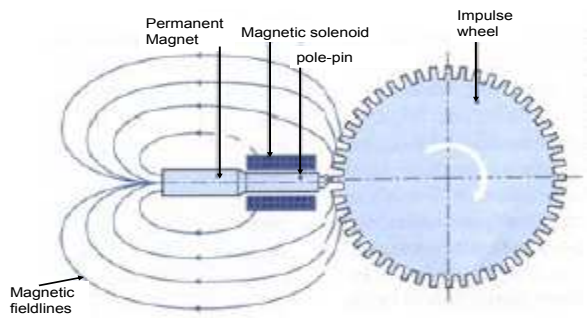

Figure 6.3: *Passive wheel speed sensors on the measuring van*

Figure (6.3) shows the components of the wheel speed sensor which have a threshold speed of about *2-5 km/h*, below which signal errors occur. The wheel sensors are switched on by the car ignition key.

Each of the four ABS wheel speed sensors is directly connected to a computer running the data collection software. The ABS sensor data collection software is a Turbo PASCAL program that generates two output files, AUTO*.DAT and TIME*.RAW. Sample outlines of the two output files are shown in Tables (6.2) and (6.3) respectively. The four ABS wheel speed sensor measurements, the derived GPS time and barometer measurements are recorded in the AUTO*.DAT file. The * indicates the file version, an integer number. When there is no barometer connected to the system, the barometer readings in column six in AUTO*.DAT file are recorded as zeros. The drive direction indicator, derived from drive direction sensor, is recorded in seventh column in this file where the ones indicate forward drive while the zeros indicate backward drive. The TIME.RAW file stores all the computations and details in the PC-time to GPS-time conversion process. These include the week-day number (Sunday is assumed to be day number zero), the UTC time, the PC-time to GPS-time conversion error and the derived GPS-time. The wheel sensors pulse signals registered by the program has a maximum range of $2^{16} - 1 = 65535$ after which the counter rolls back to zero.

Table 6.2: *Sample data of the AUTO*.DAT file*

	ABS Sensor Readings					
GPS Time	Front Right	Front Left	Back Right	Back Left	Barometer value	Direction Indicator
315982.900143	0	0	0	0	0.00	1
315983.000146	0	0	0	0	0.00	1
315983.100143	0	0	0	0	0.00	1
315983.200142	0	0	0	0	0.00	1
315983.300297	0	0	0	0	0.00	1
315983.400145	0	0	0	0	0.00	1
315983.500143	0	0	0	0	0.00	1
315983.600147	0	0	0	0	0.00	1
315983.700150	0	0	0	0	0.00	1
315983.800148	0	0	0	0	0.00	1
315983.900145	0	0	0	0	0.00	1

Table 6.3: *Sample data of the TIME*.RAW file*

PC time	Time Counters	Week day No.	GPS time	UTC time	GPS-UTC time	GPS time error	PC time error
64013.840846	17 51323 00889	3	315982	15 46 09	251968.160519	5	0
64014.840987	17 51342 52756	3	315983	15 46 10	251968.160383	5	0
64015.841126	17 51360 39090	3	315984	15 46 11	251968.160245	5	0
64016.841262	17 51378 25427	3	315985	15 46 12	251968.160108	5	0
64017.841474	17 51396 11674	3	315986	15 46 13	251968.159965	5	0
64018.841539	17 51415 63631	3	315987	15 46 14	251968.159828	5	0
64019.841679	17 51433 49964	3	315988	15 46 15	251968.159691	5	0
64020.841817	17 51451 36299	3	315989	15 46 16	251968.159554	5	0
64021.841959	17 51469 22629	3	315990	15 46 17	251968.159416	5	0
64022.842096	17 51487 08965	3	315991	15 46 18	251968.159279	5	0

6.1.3 The experiments

The field data collection was planned such that the satellites in view were at least five during the test drive while the control station *1* (Ogonda, 2003) at the top of the institute of navigation's building was used to log GPS reference data for DGPS post-processing analysis. The test drive shown in Figure (6.5) was carried out around Stuttgart city centre to collect the data used in the experiment. The equipment mounted on the institutes van was test driven for about *20 minutes* and data sampled at *100 Hz*, *10 Hz* and *1 Hz* by the MotionPak™ , ABS system and Applanix POS AV^{TM} system measurements respectively. The total test drive path is about seven kilometres. The data measured by the Applanix POS AV system is processed using Applanix POS Proc software allowing integration of the IMU measurements with the phase differential solution of the GPS positions. The solutions obtained from the POS Proc software analysis are assumed to be the actual values (true values) for all subsequent analysis.

After the synchronization of the ABS wheel speed sensors and MotionPak™ measurements with the Applanix POS-AV™ system measurements, four tests are then carried out using the field measured data. In the first test, part of the field data is used to calibrate the ABS wheel speed sensors. A loosely coupled Kalman filter process is applied in the calibration and integration of the ABS wheel speed sensors and the MotionPak™ accelerometer measurements. The second test is then performed to verify the developed discrete Kalman filter based ABS and accelerometer integration algorithms using the computer simulated data as well as the measured field data. A loosely coupled Kalman filter process is applied in the calibration and integration of the ABS wheel speed sensors and the MotionPak™ accelerometer measurements. The third experiment is to illustrate and assess the accuracy of vehicle heading estimation from the single-axis gyroscope measurements. Before the estimation of the vehicle headings, the measured angular rates from the single-axis gyro are corrected for gravitational and pitch angle effects. The last experiment is to test the algorithms developed to recover geometric road design elements, also using the simulated data before being applied on the field data. The heading differences and distance increments are used to estimate the position of the vehicle along the drive path. The generated vehicle position time series are used in the spline based curvature estimation model to derive the geometric road design parameters. A schematic view of the measurement data flow, the integration and estimation process is shown in Figure (6.4).

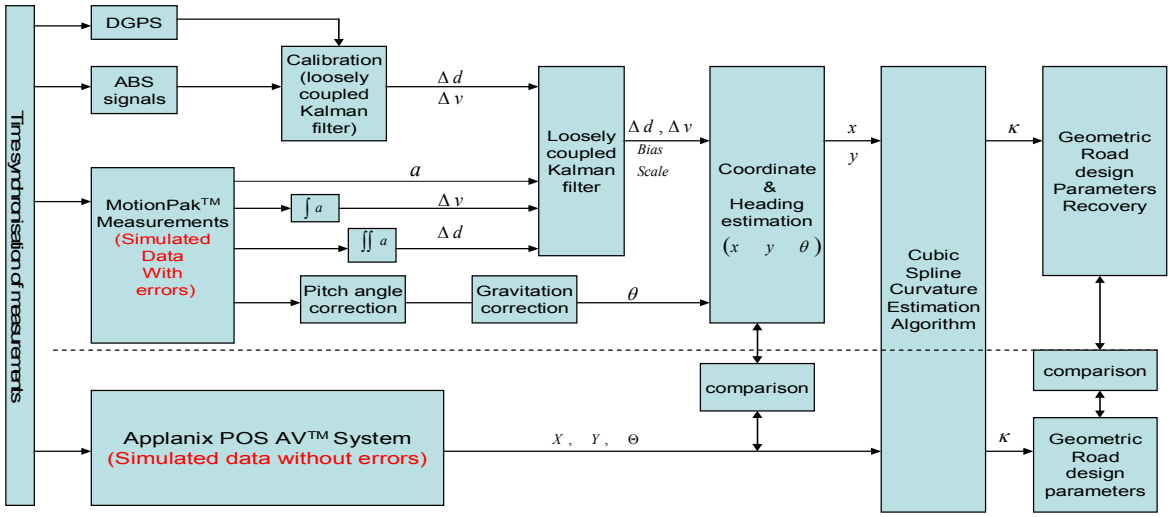

Figure 6.4*: The proposed geometric road design elements recovery model*

Figure 6.5: *Experimental drive track around Stuttgart city centre.*

6.2 Synchronisation and calibration

The accuracy of the synchronisation process of the various instruments on board the measuring vehicle platform is a factor of the vehicle speed and the sampling rates. An alternative solution is to determine the speed at which the measurement vehicle should travel during the data collection process. For example, for a vehicle travelling at *36 km/h* an accuracy of $\pm$*5 cm* can only be achieved when the synchronisation is done at better than *5 ms* for a sampling rate of *1Hz*. An accuracy of $\pm$*0.5 m* is achievable when the synchronisation is done at better than *0.05 seconds*. The synchronisation of the measurements is done in two phases. The synchronisation of ABS wheel speed sensors and the Applanix POS-AV™ system measurements at *10 Hz* is described in Section (6.2.1) while in Section (6.2.2) the synchronisation of the MotionPak™ and Applanix POS-AV™ measurements at *100 Hz* is reported.

6.2.1 ABS wheel speed sensors and Applanix POS-AV™

In the calibration of ABS and MotionPak, first a test is performed to check the time conversion process of the ABS system data collection program for the purpose of synchronisation of the data with the Applanix POS AV™ system measurements. The equipment set-up shown in Figure (6.6) is used with the other aim of the experiment being, to clear the doubt regarding the order in which the ABS wheel speed sensors were fitted on the wheels (rear, front, left, right) and the order of data logging in the output files generated by the Turbo PASCAL program used in the data collection process.

The pulses generated by the ABS wheel speed sensors are recorded in the ABS computer at a frequency of *10 Hz*. Two back wheel sensors were also connected to the Applanix POS-AV™ system and every pulse generated by the ABS sensors simultaneously recorded as event data. The Event data files are then compared to the ABS data collection program output file, AUTO*.DAT. Sample event data recorded by the Applanix system POS AV™ is shown Table (6.4) while a sample of the output file AUTO*.DAT is shown in Table (6.5). From the analysis and comparison of the measurements, it is shown that the two data sets have no significant time transformation error. The time synchronisation is therefore done one to one.

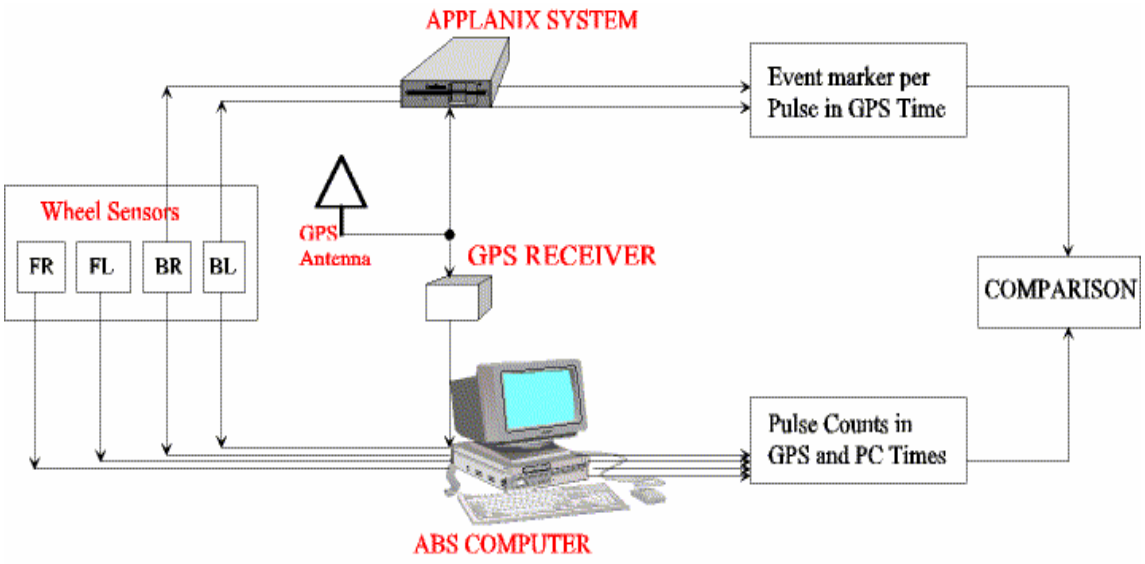

Figure 6.6: *ABS sensors time testing and synchronisation set-up*

Table 6.4: *Sample event data from Applanix POS-AV™ software*

Time_1	Event_1	Time_2	Event_2
390416.613752	1	390416.389545	1
390416.620689	2	390416.507109	2
390416.685365	3	390416.591714	3
390416.740811	4	390416.661361	4
390416.790542	5	390416.721534	5
390416.834939	6	390416.774333	6
390416.876129	7	390416.821709	7
390416.914405	8	390416.864611	8
390416.950075	9	390416.904357	9
390416.983702	10	390416.941220	10

Table 6.5: *Sample AUTO*.DAT file from the TURBO Pascal Program*

Time	FR	FL	BR	BL		
390416.300152	0	0	0	0	0.00	1
390416.400157	0	0	0	1	0.00	1
390416.500153	1	0	0	1	0.00	1
390416.600156	2	2	0	3	0.00	1
390416.700152	3	4	3	4	0.00	1
390416.800146	5	6	5	6	0.00	1
390416.900589	7	8	7	8	0.00	1
390417.000149	10	11	10	11	0.00	1
390417.100150	14	15	13	15	0.00	1
390417.200147	17	18	17	18	0.00	1

6.2.2 MotionPak™ and Applanix POS-AV™

The MotionPak has an analogue/digital (A/D) multi function converter of type ATM10 16X from national instruments. The analogue signals measured by the inertial sensors are connected to six different analogue to digital conversion channels on the card. When the IMU is integrated with a GPS or any other instrument, the converter card processor sends and receives four time signals for synchronisation purposes. As shown in Figure (6.8), the A/D converter accepts eight different analogue signal inputs within the range of $\pm 10\ V$ with a maximum sampling frequency of *100 KHz*. The individual signals are amplified through gain values. The data collection process is done in two levels. First, the output analogue data is digitalised using the A/D converter connected to the computer. The digitalised data is then read and stored using the data collection and analysis software (INTDAQ) written in C++ running in a DOS environment. Figure (6.7) shows a screen shot of the INTDAQ program that is used in real time data collection and post processing of the data. A detailed description of the data collection program is contained in Jöns (2001).

The MotionPak™ data collection program INTDAQ generates three files. The first data file usually named by user when the program is started records the time, *x*-axis angular velocity, y-axis angular velocity, *z*-axis angular velocity, *x*-axis acceleration, *y*-axis acceleration, *z*-axis acceleration and the temperature in binary format. A portion of the data file, after conversion to ASCII data format, is listed in Table (6.6). The time always start at *0.0100 second* unless the data recording is stopped and started without ending the program. The second file generated in ASCII text format is the TRATIMES.OUT, which stores the time in micro-seconds since the beginning of the recording. The last file is the MotionPak™ event data file containing the time when the pulses are generated and sent to an external equipment, in this case, the Applanix POS-AV™ system. The time T_{PC} recorded in this MotionPak™ event file is compared to the Applanix event time T_{GPS} recorded whenever an *INTDAQ* generated pulse

is received generated a linear system using Equation (6.1) and solved using least-squares technique to obtain the unknown coefficients a_0, a_1, a_2.

$$T_{GPS} = T_{PC} + a_0 + a_1 T_{PC} + a_2 T_{PC}^2 \quad (6.1)$$

The synchronization parameters obtained after the time transformation process using the least-squares technique described in Figure (6.9).

a_0 = 137242.738762742, a_1 = 0.000010249, a_2 = 0.000000000

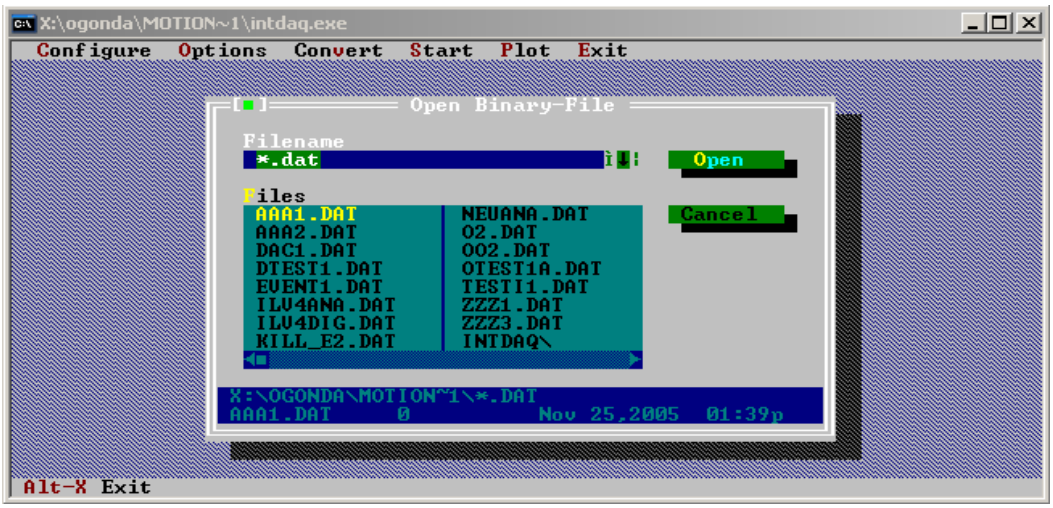

Figure 6.7: *A screen shot of the INTDAQ data collection program*

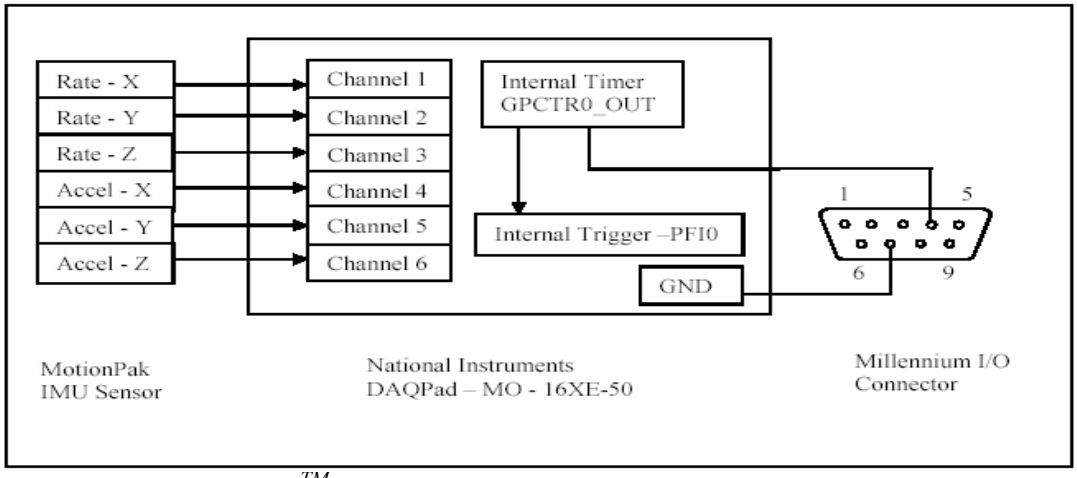

Figure 6.8: *MotionPak™ hardware connections*

Table 6.6: *Raw MotionPak™ data file generated by INTDAQ program*

t_i	w_x	w_y	w_z	a_x	a_y	a_z	T_i
0.0100	0.13173	-0.11694	-0.10717	-0.37759	0.01393	-9.89282	22.17280
0.0200	0.03415	-0.11207	-0.22896	-0.36161	0.01194	-9.89083	22.20270
0.0300	-0.12197	0.08284	-0.13153	-0.37559	0.00995	-9.88883	22.20270
0.0400	0.04879	0.05847	-0.35562	-0.36760	0.00597	-9.89680	22.20270
0.0500	0.08294	0.10233	-0.27280	-0.37959	0.01592	-9.88684	22.20270
0.0600	-0.08294	0.02436	-0.33126	-0.36960	0.00995	-9.88883	22.17280
0.0700	-0.02927	0.07796	-0.24845	-0.38159	0.01592	-9.88086	22.20270
0.0800	-0.06830	-0.05360	-0.25332	-0.37559	0.01393	-9.88883	22.20270
0.0900	0.00976	0.05847	-0.16563	-0.38958	0.01393	-9.87887	22.20270
0.1000	-0.05855	0.02436	-0.22896	-0.38558	0.01592	-9.87887	22.17280

The determined parameters are then used to transform the MotionPak™ event file and the MotionPak™ data file into GPS time. The converted MotionPak™ time is not synchronised with the Applanix POS-AV™ data time records. Therefore to estimate the MotionPak™ measured data values at the same epoch as the Applanix data time, the required MotionPak™ data value at time (t) is obtained by linear interpolation of the previous data value (x_i) and the following data value (x_{i+1}) using Equation (6.2). This enables the comparison of the MotionPak™ and the Applanix POS-AV™ data sets.

$$x = x_i + \frac{x_{i+1} - x_i}{t_{i+1} - t_i}\left(t - t_i\right) \tag{6.2}$$

The accelerations are then corrected for pitch angle effect using Equation (6.3) where g is the acceleration due gravity (approximately *9.81 m/s²*), α is the pitch angle, a is the horizontally reduced acceleration and $\hat{a}$ the measured acceleration.

$$a = \hat{a}\cos\alpha + g\sin\alpha \tag{6.3}$$

The pitch angle α is derived from the ground truth data measured using the Applanix system.

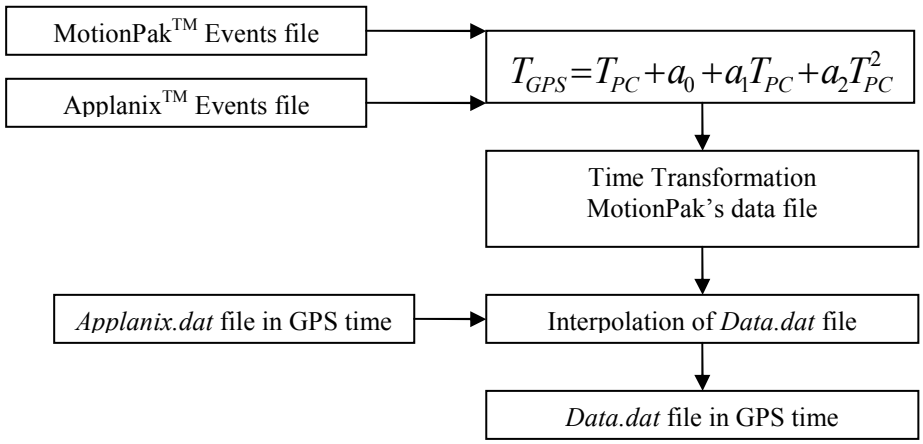

Figure 6.9: *MotionPak data file synchronization process*

6.2.3 ABS calibration

Calibration of ABS wheel speed sensors consists of the determination of the parameters required to reconstruct the wheels differential distances from which the vehicle's relative position i.e. distance increments and heading differences can be estimated. The absolute vehicle position is then be derived starting at a known configuration through discrete time integration of the wheel displacement in a process called differential odometry. The calibration process mentioned briefly in Section (3.2.2) can be carried out using least-squares technique where the parameters used in the calculation of the vehicle position from the ABS wheel sensor measurements are determined in real time or in post processing. The calibration process may also be done *geometrically* where a pre-defined trajectory is executed and processed, or *numerically* by incorporating multi-sensors through a Kalman filter algorithm. The latter approach has been employed in this synchronisation process with the Applanix POS AV™ system as the aiding. Each of the wheel speed sensors is calibrated separately generating four conversion (scale) factors but the mean rear wheel distance increments are used in subsequent computations and experiments.

An experiment was carried out to determine the number of pulses generated by the ABS sensors per wheel turn and secondly, the minimum speed at which the ABS wheel rotation sensors looses sensitivity. This speed is referred to in this report as the "*threshold speed"* of the wheel speed sensors. This Section presents the field experiments carried out to test the ABS wheel speed sensors, the formulation and data processing for the determination of the scale factor and threshold speed. The calibration process generates a scalar relating the wheel rotations to the distance travelled by the vehicle. The vehicle was driven on a straight stretch of the road allowing speed data verses individual wheel rotation signals to be developed. The analysis using Kalman filter to derive the conversion factors for converting the rotation sensor signals into distance increments follows below.

Formulation of the calibration model

From Equation (3.7), the distance increments covered by each of the left and right wheels can be estimated as shown by Equation (6.4).

$$\Delta \mathrm{d}_{\mathrm{L}} = \frac{\Delta P_L}{GR \times PPT} 2\pi R \quad and \quad \Delta \mathrm{d}_{\mathrm{R}} = \frac{\Delta P_R}{GR \times PPT} 2\pi R \tag{6.4}$$

where $\Delta d_L, \Delta d_R$ are the distance increments travelled by the left and right wheels respectively, ΔP_L, ΔP_R are the measured sensor pulse differences per time interval measured by the left and right wheel sensors respectively, R is the wheel radius, GR is the gear ratio and PPT is the number of pulses per wheel turn. Since the radius of the wheels, the gear ratio and the pulse per wheel turn may be assumed to be constants within a set of measurement, Equation (6.4) can then be rewritten for each single wheel as;

$$\Delta \mathrm{d}_{\mathrm{L}} = \Delta P_L \times SF_L \quad , \quad \Delta \mathrm{d}_{\mathrm{R}} = \Delta P_R \times SF_R \tag{6.5}$$

where $SF = \frac{2\pi R}{GR \times PPT}$ is the distance conversion (scale) factor for each of the four wheel speed sensors. The measured cumulative distance, $\hat{d}_i$ travelled by the measurement vehicle at an epoch i is then given by

$$\hat{d}_i = \hat{d}_{i-1} + \Delta \hat{d}_i + e_i \tag{6.6}$$

where $\Delta \hat{d}_i$ is the distance increment containing the normally distributed random measurement noise e_i. Inserting Equation (6.5) into (6.5) we have

$$\hat{d}_i^L = \hat{d}_{i-1}^L + \Delta P_i^L \cdot SF_i^L + e_i^L \text{ and } \hat{d}_i^R = \hat{d}_{i-1}^R + \Delta P_i^R \cdot SF_i^R + e_i^R \tag{6.7}$$

Taking subscripts RL, RR, FL, FR to denote rear-left, rear-right, front-left and front-right wheels respectively, the distances travelled by each of the four wheels at an epoch (i) are then given as;

$$\begin{aligned}
\hat{d}_i^{RL} &= \hat{d}_{i-1}^{RL} + \Delta P_i^{RL} \cdot SF_i^{RL} + e_i^{RL} \\
\hat{d}_i^{RR} &= \hat{d}_{i-1}^{RR} + \Delta P_i^{RR} \cdot SF_i^{RR} + e_i^{RR} \\
\hat{d}_i^{FL} &= \hat{d}_{i-1}^{FL} + \Delta P_i^{FL} \cdot SF_i^{FL} + e_i^{FL} \\
\hat{d}_i^{FR} &= \hat{d}_{i-1}^{FR} + \Delta P_i^{FR} \cdot SF_i^{FR} + e_i^{FR}
\end{aligned} \tag{6.8}$$

Equation (6.8) can be re-arranged in matrix form as shown in Equation (6.9).

$$\begin{bmatrix} \hat{d}^{RL} \\ \hat{d}^{RR} \\ \hat{d}^{FL} \\ \hat{d}^{FR} \end{bmatrix}_i = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \hat{d}^{RL} \\ \hat{d}^{RR} \\ \hat{d}^{FL} \\ \hat{d}^{FR} \end{bmatrix}_{i-1} + \begin{bmatrix} \Delta P^{RL} & 0 & 0 & 0 \\ 0 & \Delta P^{RR} & 0 & 0 \\ 0 & 0 & \Delta P^{FL} & 0 \\ 0 & 0 & 0 & \Delta P^{FR} \end{bmatrix}_i \begin{bmatrix} SF_{RL} \\ SF_{RR} \\ SF_{FL} \\ SF_{FR} \end{bmatrix} + \begin{bmatrix} e_d^{RL} \\ e_d^{RR} \\ e_d^{FL} \\ e_d^{FR} \end{bmatrix} \tag{6.9}$$

This is a linear model for the conversion factors $SF_{RL}, SF_{RR}, SF_{FL}, SF_{FR}$. Assuming the scale factors to be random constants due to their slow time varying characteristics, the system can therefore be solved through a simple Gauß-Markov process of the first order given in matrix form as shown in Equation (6.10).

$$\begin{bmatrix} SF_{RL} \\ SF_{RR} \\ SF_{FL} \\ SF_{FR} \end{bmatrix}_i = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} SF_{RL} \\ SF_{RR} \\ SF_{FL} \\ SF_{FR} \end{bmatrix}_{i-1} + \begin{bmatrix} e_S^{RL} \\ e_S^{RR} \\ e_S^{FL} \\ e_S^{FR} \end{bmatrix} \tag{6.10}$$

Combining Equations (6.9) and (6.10) results in Equation (6.11) which is an augmented linear model of the form $x_i = Ax_{i-1} + B\varepsilon_i$.

$$\begin{bmatrix} \hat{d}^{RL} \\ \hat{d}^{RR} \\ \hat{d}^{FL} \\ \hat{d}^{FR} \\ SF_{RL} \\ SF_{RR} \\ SF_{FL} \\ SF_{FR} \end{bmatrix}_i = \begin{bmatrix} 1 & 0 & 0 & 0 & \Delta P_i^{RL} & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & \Delta P_i^{RR} & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & \Delta P_i^{FL} & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & \Delta P_i^{FR} \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \hat{d}^{RL} \\ \hat{d}^{RR} \\ \hat{d}^{FL} \\ \hat{d}^{FR} \\ SF_{RL} \\ SF_{RR} \\ SF_{FL} \\ SF_{FR} \end{bmatrix}_{i-1} + \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} e_d^{RL} \\ e_d^{RR} \\ e_d^{FL} \\ e_d^{FR} \\ e_S^{RL} \\ e_S^{RR} \\ e_S^{FL} \\ e_S^{FR} \end{bmatrix}_i \tag{6.11}$$

Measurement updates using the GPS solution

The measurement updates from the Applanix POS-AV™ system used in the calibration process are derived as follows. Assuming that Figure (6.10) represents the wheels of a vehicle moving in a circular curve centred at point O at two epochs $(i, i+1)$, the relationship of the ABS derived distance increments, $\Delta\tilde{d}$ and the Applanix POS-AV™ derived increments, Δd is derived from the geometry of the figure as follow. For any given time interval, Δt

$$C = \frac{\Delta d}{\Delta \alpha} \quad \Rightarrow \quad \Delta d = C \cdot \Delta \alpha \tag{6.12}$$

$$F = [C^2 + L^2]^{1/2} = \left[\left(\frac{\Delta d}{\Delta \alpha} \right)^2 + L^2 \right]^{1/2} \tag{6.13}$$

where Δd is the distance increments, $\Delta\alpha$ is the heading differences measured by the Applanix POS-AV™ system, C is the radius of circular curve covered by the rear wheels, F is the radius of circular curve covered by the front wheels and L the length between the front and rear wheels. Taking B to be the base length between the wheels and using the relation in Equation (6.5), then the distance increments travelled by the rear wheel ABS sensors are given as in Equation (6.14) and (6.15).

$$\Delta\hat{d}^{RL} = \Delta P^{RL} \cdot SF_{RL} = \left(C + \tfrac{B}{2}\right) \cdot \Delta\alpha = C \cdot \Delta\alpha + \tfrac{B}{2} \cdot \Delta\alpha = \Delta d + \tfrac{B}{2} \cdot \Delta\alpha \tag{6.14}$$

$$\Delta\hat{d}^{RR} = \Delta P^{RR} \cdot SF_{RR} = \left(C - \tfrac{B}{2}\right) \cdot \Delta\alpha = C \cdot \Delta\alpha - \tfrac{B}{2} \cdot \Delta\alpha = \Delta d - \tfrac{B}{2} \cdot \Delta\alpha \tag{6.15}$$

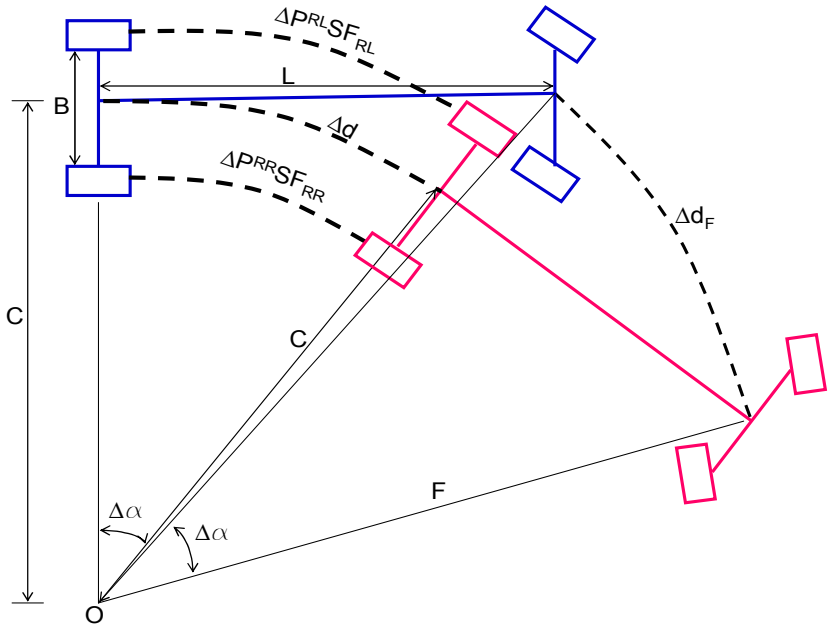

Figure 6.10: *Differential odometry*

And similarly the distance increments travelled by the front wheel is given in Equations (6.16) and (6.17).

$$\Delta\hat{d}^{FL} = \Delta P^{FL} \cdot SF_{FL} = \left(F + \tfrac{B}{2}\right) \cdot \Delta\alpha = \left(\left[C^2 + L^2\right]^{1/2} + \tfrac{B}{2}\right) \cdot \Delta\alpha = \left[\Delta d^2 + L^2 \cdot \Delta\alpha^2\right]^{1/2} + \tfrac{B}{2}\Delta\alpha \tag{6.16}$$

$$\Delta\hat{d}^{FR} = \Delta P^{FR} \cdot SF_{FR} = \left(F - \tfrac{B}{2}\right) \cdot \Delta\alpha = \left(\left[C^2 + L^2\right]^{1/2} - \tfrac{B}{2}\right) \cdot \Delta\alpha = \left[\Delta d^2 + L^2 \cdot \Delta\alpha^2\right]^{1/2} - \tfrac{B}{2}\Delta\alpha \tag{6.17}$$

which means that;

$$\Delta d_i + \tfrac{B}{2}\Delta\alpha_i = SF_{RL}\Delta P_i^{RL}$$

$$\Delta d_i - \tfrac{B}{2}\Delta\alpha_i = SF_{RR}\Delta P_i^{RR}$$

$$\left[\Delta d_i^2 + L^2 \cdot \Delta\alpha_i^2\right]^{1/2} + \tfrac{B}{2}\Delta\alpha_i = SF_{FL}\Delta P_i^{FL} \tag{6.18a}$$

$$\left[\Delta d_i^2 + L^2 \cdot \Delta\alpha_i^2\right]^{1/2} - \tfrac{B}{2}\Delta\alpha_i = SF_{FR}\Delta P_i^{FR}$$

Since the quantities in the left hand side of Equation (6.18a) are known from the GPS solution, the ABS wheel sensors can therefore be calibrated using the GPS solutions by solving the matrix given in Equation (6.18) below.

$$\begin{bmatrix} \Delta d + \frac{B}{2}\cdot\Delta\alpha \\ \Delta d - \frac{B}{2}\cdot\Delta\alpha \\ \left(\Delta d^2 + L^2\cdot\Delta\alpha^2\right)^{1/2} + \frac{B}{2}\cdot\Delta\alpha \\ \left(\Delta d^2 + L^2\cdot\Delta\alpha^2\right)^{1/2} - \frac{B}{2}\cdot\Delta\alpha \end{bmatrix}_i = \begin{bmatrix} \Delta P^{RL} & 0 & 0 & 0 \\ 0 & \Delta P^{RR} & 0 & 0 \\ 0 & 0 & \Delta P^{FL} & 0 \\ 0 & 0 & 0 & \Delta P^{FR} \end{bmatrix}_i \begin{bmatrix} SF_{RL} \\ SF_{RR} \\ SF_{FL} \\ SF_{FR} \end{bmatrix} + \begin{bmatrix} e_S^{RL} \\ e_S^{RR} \\ e_S^{FL} \\ e_S^{FR} \end{bmatrix}_i \quad (6.18)$$

The results obtained from the calibration of the ABS using the MATLAB codes of the above model formulated above are shown in Figures (6.11) and (6.12). A summary of the results is tabulated in Table (6.7) below. From the table, the rear left scale factor has the lowest standard deviation of ± 0.00148261. The distance increments computed using the estimated scale factors for each wheel are shown in Figure (6.13). The figures illustrate that in terms errors in distance increments, obtained by comparing to the Applanix POS AV™ system derived distances, the back left wheel speed sensor shows relatively better stability with a mean distance increments error of *0.0000217 m* and a standard deviation of ± 0.023838 *m*. Estimation of the cumulative distances using the rear left sensor scale factor of *0.036849552* also results in the lowest maximum cumulative error of about *10m*. This may be attributed, but not necessarily true, to the fact that the equipment assembly was mounted on the left side of the vehicle and therefore the left wheel experienced less perturbation during the test drive due to the weight differences. From the mean standard deviation of distance difference estimated by the rear wheels, the accuracy of the distance difference can be concluded to be $\pm 0.02m$, as estimated from the mean of the rear wheel speed sensor scale factors.

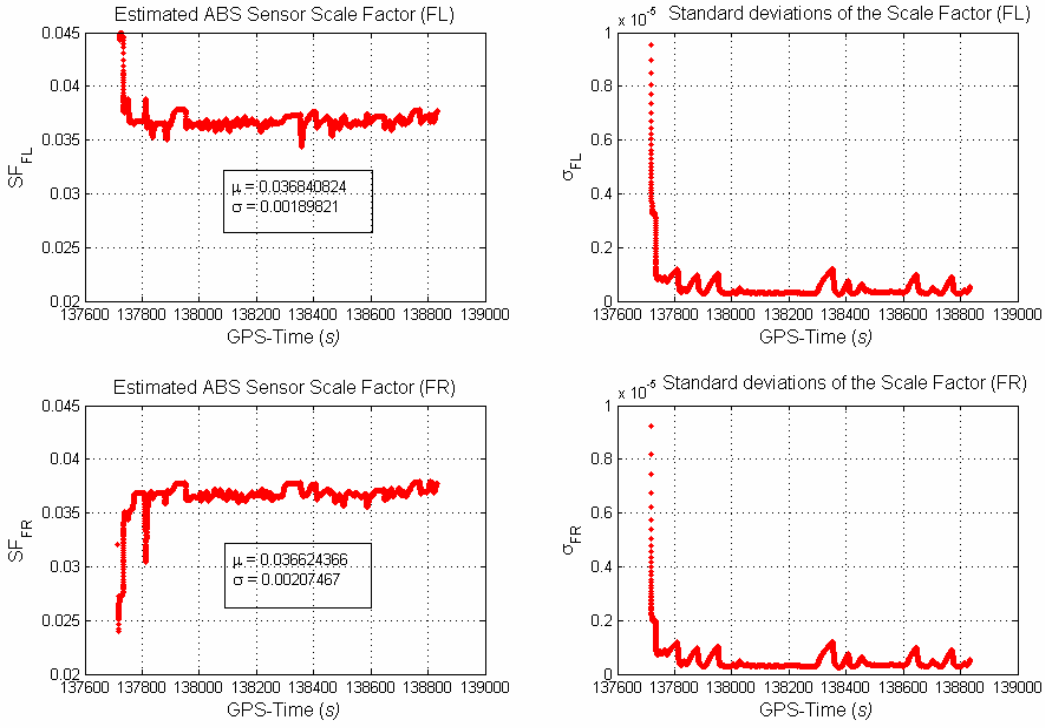

Figure 6.11: *Estimation of ABS wheel speed sensors conversion factors (front wheels)*

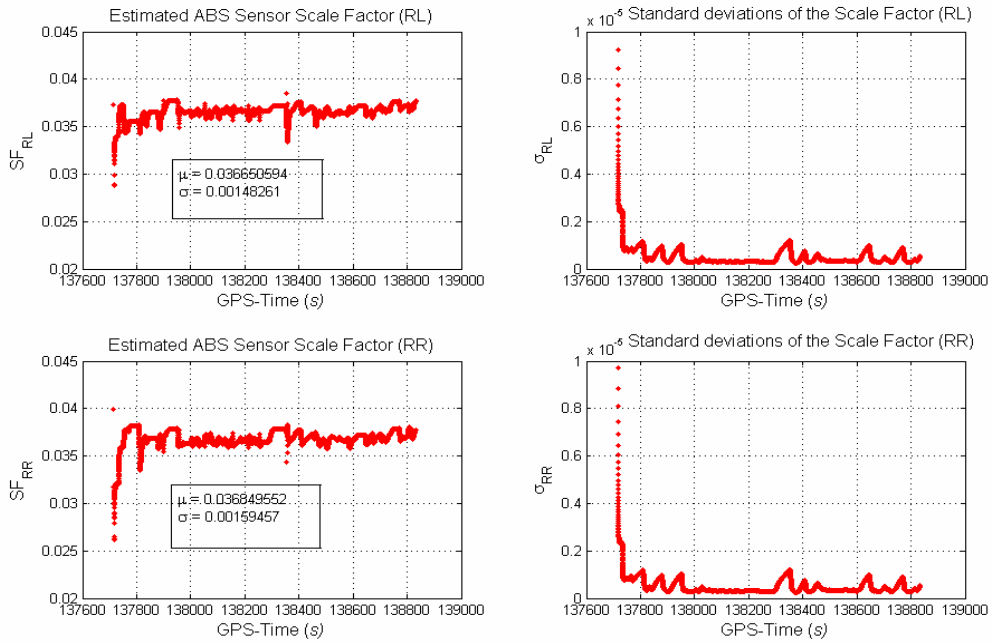

Figure 6.12: *Estimation of ABS wheel speed sensors conversion factors (rear wheels)*

Table 6.7: *Results of the ABS sensors calibration*

	Scale Factor	Standard Deviation
Rear Left (RL)	0.036650594	0.00148261
Rear Right (RR)	0.036849552	0.00159457
Front Left (FL)	0.036840824	0.00189821
Front Right (FR)	0.036624366	0.00207467

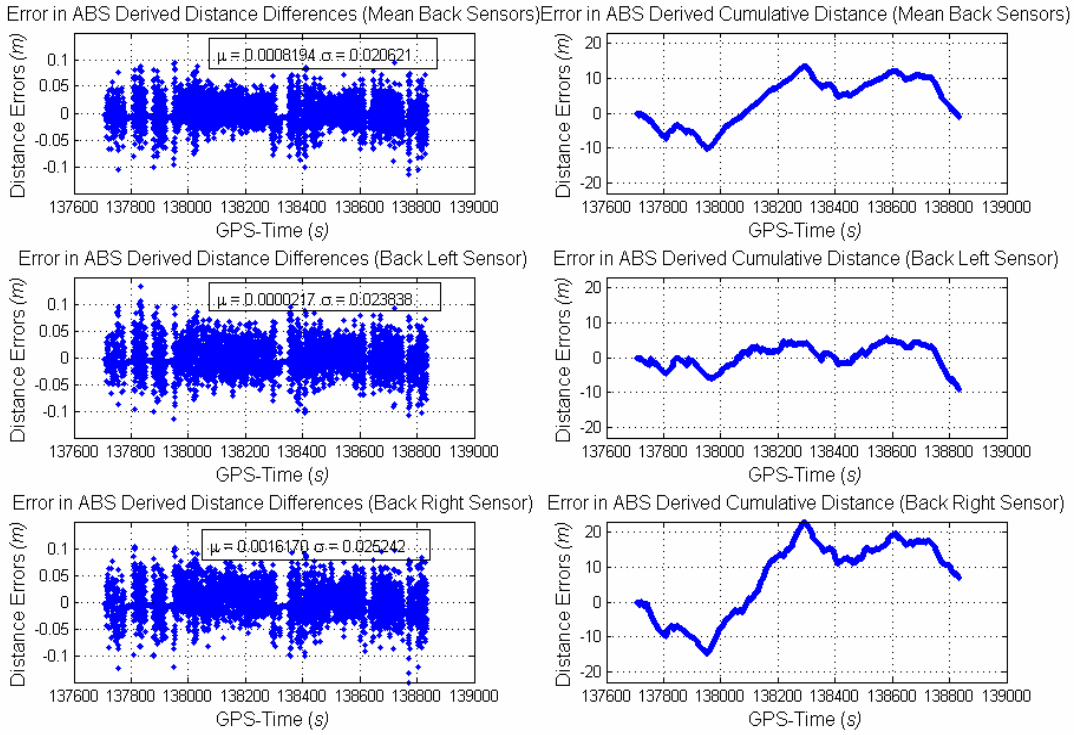

Figure 6.13: *Errors in ABS derived distance differences and cumulative distances*

6.2.3.1 Estimation of the threshold speed

As mentioned earlier, the wheel speed sensors that are based on variable reluctance technique are dependent on the speed of rotation of the wheels i.e. there exists a threshold speed below which the signals generated are too weak to be sensed hence the signal truncation. The ABS was then used in an experiment carried out at the University of Hohenheim in Plieningen, where the vehicle was driven in a straight line but at varied speeds at an open site with a clear view of the sky. From the analysis one sees that if no pulse was measured within the time interval, then the sensitivity can be said to correspond to at least *1* pulse per turn. Considering the data sampling rate of *10 Hz* by our program, then;

$$\begin{aligned} 1 \quad \text{pulse} / 0.1 \,\text{second} &\Rightarrow 10 \,\text{pulses} / \text{second} \\ &\Rightarrow 10 \times SF \\ &\Rightarrow 10 \times 0.03667 \; m/s \\ &\Rightarrow 1.32 \; Km/h \end{aligned} \tag{6.19}$$

That is, if the vehicle is moving at velocities less than *1.32Km/h*, the ABS sensors data collection program will register no measurements. This compares well with the manufacturer defined threshold speed of *2-5 km/h.* The threshold speed and is dependent on the number of ABS sensors teeth (pulse per turn) and the vehicle wheel radius.

6.3 Integration of ABS and accelerometer

After the determination of the threshold speed of the ABS wheel speed sensors, the data integration is performed by bridging the ABS wheel speed sensor measurement signal gaps with the acceleration measurements through estimation and removal of accelerometer biases and scale factor errors i.e. during the periods in which the measurement vehicle was moving at speeds lower than the *threshold speed*.

6.3.1 Mathematical formulation of the integration model

In this section, the detailed Kalman filter algorithm designed to integrate the ABS wheel speed sensor measurements and the single-axis MEMS accelerometer measurements is presented. The integrated solution is subsequently used to estimate the distance increments travelled by the vehicle wheels. The filter is based on enhancing the distance increments estimation by using the ABS measurements to estimate the accelerometer bias and scale factor errors when the vehicle speed is above the threshold speed. When the vehicle speed is lower than the threshold speed, the estimated bias and scale factor errors are used to correct the accelerometer measurements which are then integrated once to obtain the vehicle speed and twice to obtain the travelled distance (*cf.* Figure 6.14). Adopting the single axis accelerations, $\tilde{a}_x$ expressed in Equation (3.6) and assuming the cross-axis coupling factors are insignificant, then if the accelerometer bias and scale factor error are known and the gravitational correction has been applied, the equation is modified to estimate the velocity difference as shown in Equation (6.20).

$$\partial v^m(t_i) = \bar{a}(t_i, t_{i+1}) \cdot (t_i - t_{i-1}) + B \cdot (t_i - t_{i-1}) + S \cdot \partial v^m(t_i) + \varepsilon_v \tag{6.20}$$

where B is the accelerometer bias, S is the accelerometer scale factor error, $\bar{a}$ is the mean acceleration between two successive epochs and ε is random white noise. Through integration, the cumulative distances $\hat{d}$ and velocities $\hat{v}$ are obtain as

$$\tilde{d}(t_i) = \tilde{d}(t_{i-1}) + \tilde{v}(t_{i-1}) \cdot (t_i - t_{i-1}) + \tfrac{1}{2} \partial v^m(t_i) \cdot (t_i - t_{i-1}) + \tfrac{1}{2} \varepsilon_v \cdot (t_i - t_{i-1}) \quad (6.21)$$

$$\tilde{v}(t_i) = \tilde{v}(t_{i-1}) + \partial v^m(t_i) + \varepsilon_v \quad (6.22)$$

where $(\sim)$ indicates variables that are contaminated by random noise. From the basic equations of motion, the cumulative distance d, is also given as

$$d = ut + \tfrac{1}{2} at^2 \quad (6.23)$$

where u is the initial velocity, a is the acceleration and t is the time. Differentiating the equation above with respect to time once we obtain the equation for velocity as

$$v = u + at \quad (6.24)$$

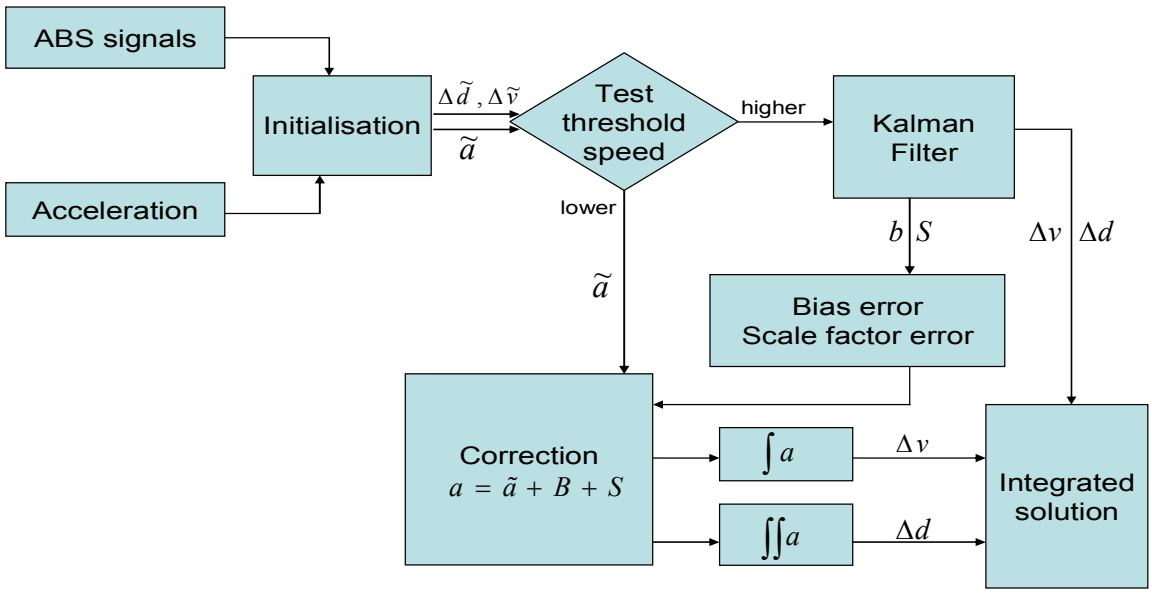

Figure 6.14: *Integration f ABS and accelerometer*

Therefore the true cumulative distance and velocity at time t_i can (theoretically) be computed from Equation (6.20) as follows

$$d(t_i) = d(t_{i-1}) + v(t_{i-1}) \cdot (t_i - t_{i-1}) + \tfrac{1}{2}[\partial v^m(t_i) - B \cdot (t_i - t_{i-1}) - S \cdot \partial v^m(t_i)] \cdot (t_i - t_{i-1}) + \tfrac{1}{2} \varepsilon_v \cdot (t_i - t_{i-1}) \quad (6.25)$$

$$v(t_i) = v(t_{i-1}) + [\partial v^m(t_i) - B \cdot (t_i - t_{i-1}) - S \cdot \partial v^m(t_i)] + \varepsilon_v \quad (6.26)$$

Taking the differences between Equations (6.21) and (6.25) and Equations (6.22) and (6.26) one obtains

$$\tilde{d}(t_i) - d(t_i) = \tilde{d}(t_{i-1}) - d(t_{i-1}) + [\tilde{v}(t_{i-1}) - v(t_{i-1})] \cdot (t_i - t_{i-1}) + \tfrac{1}{2}[B \cdot (t_i - t_{i-1}) - S \cdot \partial v^m(t_i)] \cdot (t_i - t_{i-1}) + \tfrac{1}{2} \varepsilon_v \cdot (t_i - t_{i-1}) \quad (6.27)$$

$$\tilde{v}(t_i) - v(t_i) = \tilde{v}(t_{i-1}) - v(t_{i-1}) + [B \cdot (t_i - t_{i-1}) - S \cdot \partial v^m(t_i)] + \varepsilon_v \quad (6.28)$$

$\Delta d(t_i)$ and $\Delta v(t_i)$ are defined as

$$\Delta d(t_i) = \tilde{d}(t_i) - d(t_i) \tag{6.29}$$

$$\Delta v(t_i) = \tilde{v}(t_i) - v(t_i) \tag{6.30}$$

Inserting $\Delta d(t_i)$ and $\Delta v(t_i)$ as defined in Equations (6.29) and (6.30) above, Equations (6.27) and (6.28) can then be rewritten as

$$\Delta d(t_i) = \Delta d(t_{i-1}) + \Delta v(t_{i-1}) \cdot (t_i - t_{i-1}) + \tfrac{1}{2}[B \cdot (t_i - t_{i-1}) - S \cdot \partial v^m(t_i)] \cdot (t_i - t_{i-1}) + \tfrac{1}{2}\varepsilon_v \cdot (t_i - t_{i-1}) \tag{6.31}$$

$$\Delta v(t_i) = \Delta v(t_{i-1}) + [B \cdot (t_i - t_{i-1}) - S \cdot \partial v^m(t_i)] + \varepsilon_v \tag{6.32}$$

Equations (6.31) and (6.32) are re-arranged in matrix form as

$$\begin{bmatrix} \Delta d \\ \Delta v \end{bmatrix}_{t_i} = \begin{bmatrix} 1 & t_i - t_{i-1} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \Delta d \\ \Delta v \end{bmatrix}_{t_{i-1}} + \begin{bmatrix} \frac{(t_i - t_{i-1})^2}{2} & \frac{\partial v^m(t_i) \cdot (t_i - t_{i-1})}{2} \\ (t_i - t_{i-1}) & \partial v^m(t_i) \end{bmatrix} \begin{bmatrix} B \\ S \end{bmatrix} + \begin{bmatrix} \frac{(t_i - t_{i-1})}{2} \\ 1 \end{bmatrix} \begin{bmatrix} \varepsilon_v \end{bmatrix} \tag{6.33}$$

Equation (6.33) is a linear model for the accelerometer bias B and accelerometer scale factor error S. If one assumes a random process for the slowly changing B and S then a solution obtained using, for example, a simple Gauß-Markov process of the first order is given as follows

$$B(t_i) = B_B B(t_{i-1}) + \varepsilon_B \tag{6.34}$$

$$S(t_i) = B_S S(t_{i-1}) + \varepsilon_S \tag{6.35}$$

And in matrix form

$$\begin{bmatrix} B \\ S \end{bmatrix}_{t_i} = \begin{bmatrix} B_B & 0 \\ 0 & B_S \end{bmatrix} \begin{bmatrix} B \\ S \end{bmatrix}_{t_{i-1}} + \begin{bmatrix} \varepsilon_B \\ \varepsilon_S \end{bmatrix} \tag{6.36}$$

Combining Equations (6.33) and (6.36) results in

$$\begin{bmatrix} \Delta d \\ \Delta v \\ B \\ S \end{bmatrix}_{t_i} = \begin{bmatrix} 1 & t_i - t_{i-1} & B_B \frac{(t_i - t_{i-1})^2}{2} & B_S \frac{\partial v^m(t_i) \cdot (t_i - t_{i-1})}{2} \\ 0 & 1 & B_B(t_i - t_{i-1}) & B_S \partial v^m(t_i) \\ 0 & 0 & B_B & 0 \\ 0 & 0 & 0 & B_S \end{bmatrix} \begin{bmatrix} \Delta d \\ \Delta v \\ B \\ S \end{bmatrix}_{t_{i-1}} + \begin{bmatrix} \frac{(t_i - t_{i-1})}{2} & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \varepsilon_v \\ \varepsilon_B \\ \varepsilon_S \end{bmatrix} \tag{6.37}$$

Equation (6.37) is an augmented linear model of the form

$$x_i = A x_{i-1} + B \cdot \varepsilon \tag{6.38}$$

Given initial values

$$\begin{aligned} &\tilde{d}(t_0) = d(t_0) + \Delta d(t_0), \sigma_d^2(t_0) \\ &\tilde{v}(t_0) = v(t_0) + \Delta v(t_0), \sigma_v^2(t_0) \\ &B(t_0), S(t_0), \sigma_B^2(t_0), \sigma_S^2(t_0), \sigma_v^2, \sigma_B^2, \sigma_S^2 \end{aligned} \tag{6.39}$$

$$x_0 = \begin{bmatrix} \Delta d(t_0) \\ \Delta v(t_0) \\ B(t_0) \\ S(t_0) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \tag{6.40}$$

$$\Sigma_{A_0} = \begin{bmatrix} \sigma_d^2(t_0) & 0 & 0 & 0 \\ 0 & \sigma_v^2(t_0) & 0 & 0 \\ 0 & 0 & \sigma_B^2(t_0) & 0 \\ 0 & 0 & 0 & \sigma_S^2(t_0) \end{bmatrix} \tag{6.41}$$

$$\Sigma_B = \begin{bmatrix} \sigma_v^2 & 0 & 0 \\ 0 & \sigma_B^2 & 0 \\ 0 & 0 & \sigma_S^2 \end{bmatrix} \tag{6.42}$$

where x_0 is the initial state vector estimate, Σ_{A_0} is the initial state vector covariance and Σ_B is the initial process noise covariance. If the updates are not available, the covariance Σ_A will cumulate in time, until another update is available. The above formulation was implemented using the Kalman filtering code listed below (Kleusberg, 2004). Prediction of the state vector

$$\partial x_n(-) = F_n \cdot \partial x_{n-1}(+) \tag{6.43}$$

Prediction of the error covariance matrix

$$Cx_n(-) = F_n \cdot Cx_{n-1}(+) F_n^T + G \cdot Cw_n \cdot G^T \tag{6.44}$$

Computation of the Kalman gain factor

$$K_n = Cx_n(-) \cdot H^T \cdot (H \cdot Cx_n(-) \cdot H^T + Cv_n)^{-1} \tag{6.45}$$

Updating the state vector with the measurements

$$\partial x_n(+) = \partial x_n(-) + K_n \cdot \left[\partial y_n - H \cdot \partial x_n(-) \right] \tag{6.46}$$

Updating the error covariance matrix

$$Cx_n(+) = Cx_n(-) - K_n \cdot H \cdot Cx_n(-) \tag{6.47}$$

The accelerometer measurement updates from the ABS derived cumulative distances and velocities, we have the following Equations;

$$D^m(T_i) = D(T_i) + \varepsilon_A \quad , \quad \sigma_A^2 \tag{6.48}$$

but we know that

$$D(T_i) = \tilde{D}(T_i) - \Delta d(T_i) + \varepsilon_D \tag{6.49}$$

Therefore

$$\tilde{D}(T_i) - D^m(T_i) = \Delta d(T_i) + \varepsilon_D \tag{6.50}$$

Similarly

$$\tilde{V}(T_i) - V^m(T_i) = \Delta v(T_i) + \varepsilon_V \tag{6.51}$$

Writing Equations (6.50) and (6.51) in matrix form

$$\begin{bmatrix} \tilde{D}(T_i) - D^m(T_i) \\ \tilde{V}(T_i) - V^m(T_i) \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} \Delta d(T_i) \\ \Delta v(T_i) \\ B \\ S \end{bmatrix} + \begin{bmatrix} \varepsilon_D \\ \varepsilon_V \end{bmatrix} \tag{6.52}$$

$$\begin{bmatrix} \tilde{D}(T_i) - D^m(T_i) \\ \tilde{V}(T_i) - V^m(T_i) \end{bmatrix} = \begin{bmatrix} 1 & 0 & \frac{\Delta t^2}{2} & \partial v^m(T_i) \cdot \left(\frac{\Delta t^2}{2} \right) \\ 0 & 1 & \Delta t & \partial v^m(T_i) \end{bmatrix} \begin{bmatrix} \Delta d(T_i) \\ \Delta v(T_i) \\ B \\ S \end{bmatrix} + \begin{bmatrix} \varepsilon_D \\ \varepsilon_V \end{bmatrix} \tag{6.53}$$

where $\tilde{D}(T_i)$ are the MotionPak™ derived cumulative distances and $D^m(T_i)$ are the ABS derived cumulative distances at epochs (T_i). Equation (6.53) is of the form

$$y_i = Ax_i + \varepsilon \tag{6.54}$$

The corrected distances and velocities ($\tilde{d}$, $\tilde{v}$) are used to generate a time series of biases and scale factors errors $B(T_i)$, $S(T_i)$.

6.3.2 Data Simulation

A Kalman filter code modelled in Section (6.3.2) and written in MATLAB programming language is used to integrate the ABS and accelerometer measurements. Before the MotionPak™ accelerometer and ABS wheel speed sensor measurements could be integrated, an investigation is carried out to determine the effects of accelerometer and ABS measurement noise on the estimated accelerometer bias and scale factor error.

The comprehensive simulation tests in this section are performed to verify the formulations of the developed algorithms used to integrate the ABS wheel speed sensors, the single-axis accelerometer and GPS data sets. Mean accelerations data analogous to x-axis MotionPak™ accelerometer measurements are simulated at *100 Hz* as per Equation (6.57). The mean accelerometer data are then integrated once with respect to time to obtain velocities analogous to Applanix POS AV™ velocities, v and twice to obtain distance analogous to Applanix derived cumulative distances, d at *100 Hz* using Equations (6.55) and (6.56) respectively. Known bias, scale factor error and noise were then added to the simulated accelerometer measurements and then integrated once and twice to obtain MotionPak derived velocities and cumulative distance. The simulation of the cumulative distance, velocity and acceleration are shown in Figures (6.16), (6.17) and (6.18). The discrete Kalman filter code is then used to integrate both the simulated and measured data as shown in the data flow process in Figure (6.15) below.

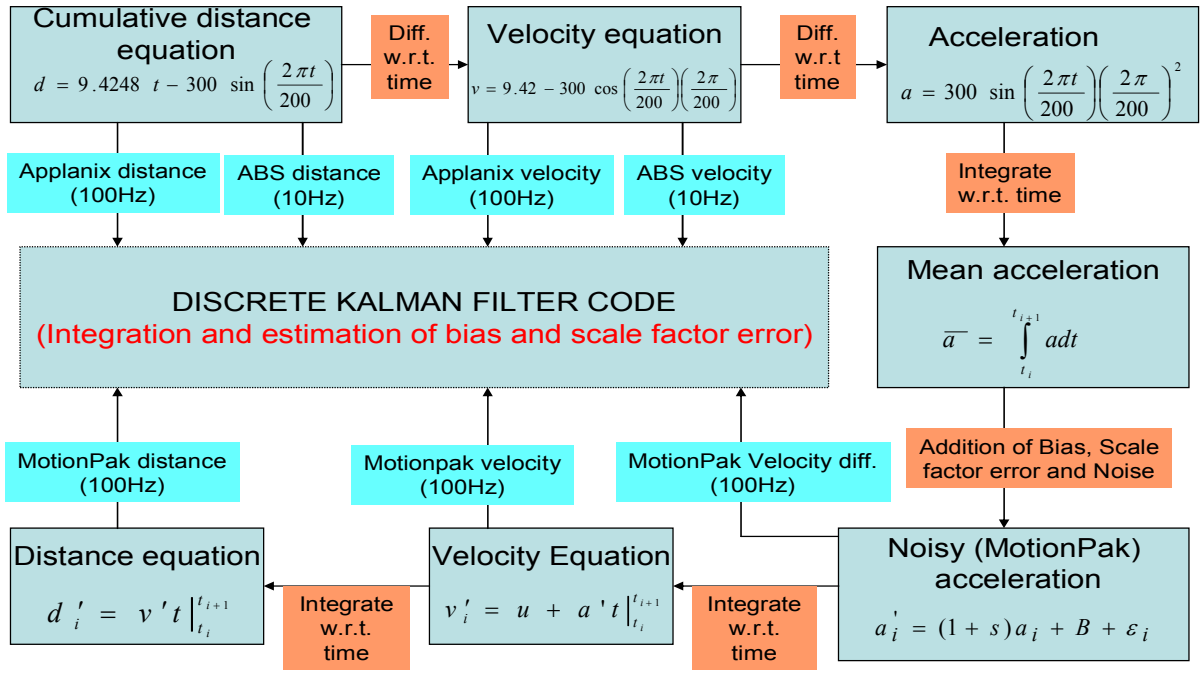

Figure 6.15: *Data simulation and integration process for testing the Kalman filter code*

$$d = 9.4248t - 300\sin\left(\frac{2\pi t}{200}\right) \tag{6.55}$$

$$v = 9.4248 - 300\cos\left(\frac{2\pi\Delta t}{200}\right)\left(\frac{2\pi}{200}\right) \tag{6.56}$$

$$a = 300\sin\left(\frac{2\pi t}{200}\right)\left(\frac{2\pi}{200}\right)^2 \tag{6.57}$$

where a is the acceleration and Δt the time interval. A bias $B = \pm 0.02ms^{-2}$ and a scale factor error $S = 3.5\%$ are introduced into the simulated mean acceleration values while Gaussian white noise (ε) of zero mean and standard deviation of ± 0.025 are added to the simulated distances and velocities. The simulated distances and velocities are assumed analogous to ABS derived distances and velocities while the noisy (contaminated) accelerometer values assumed to be analogous to MotionPak™ derived accelerations were integrated once to obtain MotionPak™ derived velocities and integrated twice to obtain MotionPak™ derived cumulative distances as indicated in Equations (6.58), (6.59) and (6.60) below.

$$a_i' = (1+s)a_i + B + \varepsilon_i \tag{6.58}$$

$$v_i = u + \sum a'\Delta t\Big|_{t_i}^{t_{i+1}} \tag{6.59}$$

$$d_i = \sum v\Delta t\Big|_{t_i}^{t_{i+1}} \tag{6.60}$$

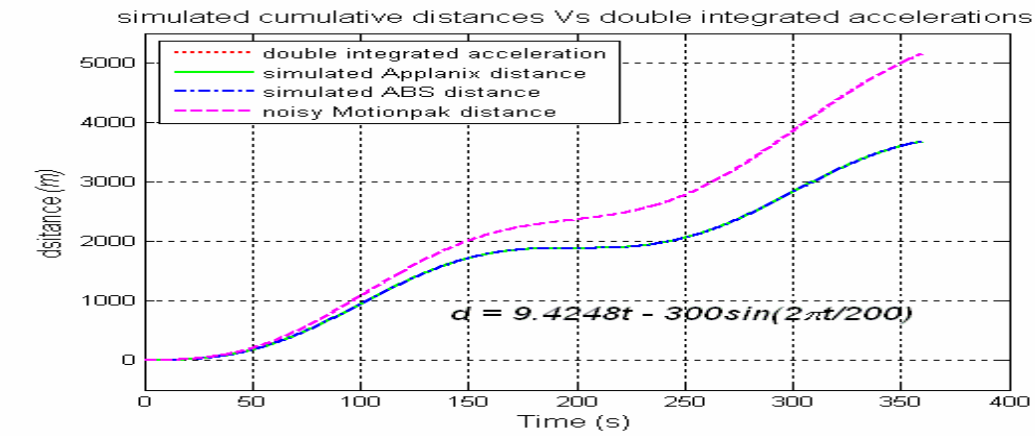

Figure 6.16: *Simulated cumulative distances*

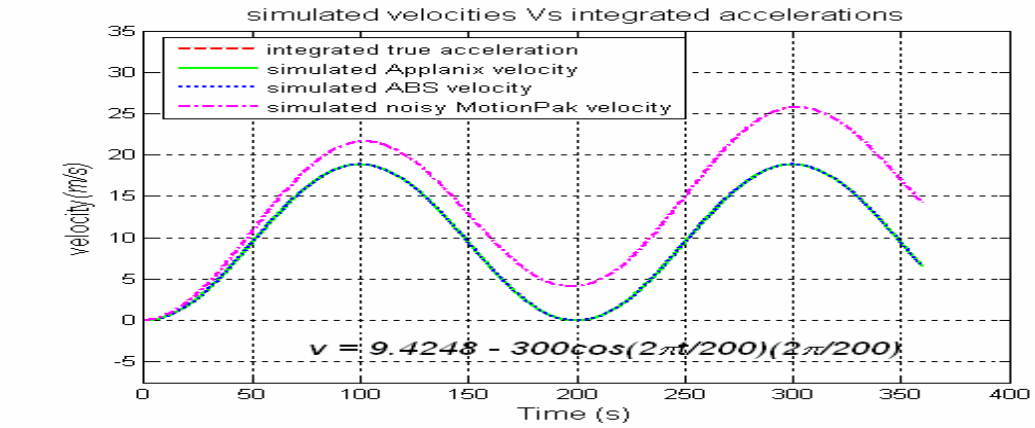

Figure 6.17: *Simulated velocities*

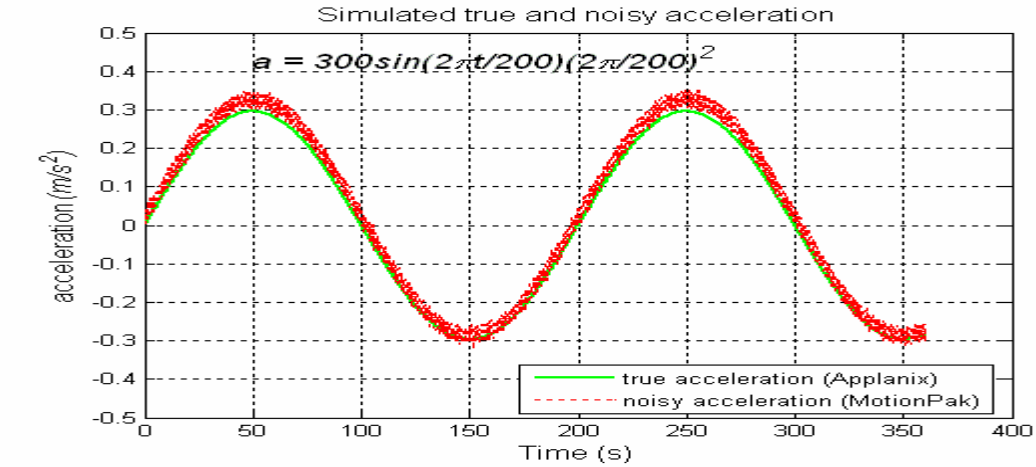

Figure 6.18: *Simulated acceleration values*

Using the different settings for the bias, scale factor error and noise as shown in Table (6.8) and (6.9), the results described in section (6.3.2.1) were realised.

Table 6.8: *Parameter settings used in testing the effects of accelerometer measurement noise*

	Test 1A	*Test 1B*	*Test 1C*
Accelerometer Bias	0.020 m/s^2	0.020 m/s^2	0.020 m/s^2
Accelerometer scale factor error	3.5%	3.5%	3.5%
Accelerometer Measurement Noise (m/s^2)	$N \sim (0, 0.000625)$	$N \sim (0, 0.00000625)$	$N \sim (0, 0)$
Initial bias-covariance	0.64	0.64	0.64
ABS Measurement Noise (m) (m/s)	Distance: $N \sim (0, 0.0185^2)$ Velocity: $N \sim (0, 0.185^2)$	Distance: $N \sim (0, 0.0185^2)$ Velocity: $N \sim (0, 0.185^2)$	Distance: $N \sim (0, 0.0185^2)$ Velocity: $N \sim (0, 0.185^2)$
Initial scale factor-covariance	0.64	0.64	0.64

6.3.2.1 Simulated data integration results

Kalman filter is widely applied as the core integration algorithm for GPS/INS schemes in many navigation solutions despite the following disadvantages having been cited. The Kalman filter solution depends on the model, prior knowledge and linearization. Simulated data is used to analyse the dependency of the integration process on the model, prior knowledge and noise levels in the measurements.

This section therefore contains the results of the tests performed to determine the influence of the accelerometer and the ABS wheel sensors' measurement noise on the estimation of the bias, scale factor errors and he integrated distance solution. A bias of *0.02 m/s^2* and the scale factor error of *3.5%* are introduced in the simulated acceleration measurements. For the first set of tests, the standard deviation of the random noise in distance increments derived from the wheel speed sensor ($\sigma = \pm 0.0185m$) is held constant while the standard deviations of the random noise in the simulated accelerometer measurements are set to $\sigma = \pm 0.025m/s^2$ in *Test 1A*, $\sigma = \pm 0.0025m/s^2$ in *Test 1B* and $\sigma = \pm 0.0m/s^2$ in *Test 1C* as summarised in Table (6.8). Using an initial state vector variance of *0.8 m/s* in speed and *8.8 m* in distance, the estimated accelerometer bias and scale factor errors are shown in Figures (6.19), (6.20) and (6.21). From the results, the standard deviation of the estimated velocities and distances increments are about $\sigma = \pm 0.01m/s$ and $\sigma = \pm 0.002m$ in all the three tests. One may conclude that reducing the accelerometer noise to millimetre level or even the elimination of the accelerometer noise altogether seems not to improve or worsen the estimated velocities and distances. The accelerometer noise can therefore be said to have minimal effect in relation to the wheel speed sensors measurements noise in the estimation of the bias, scale factor error and distances.

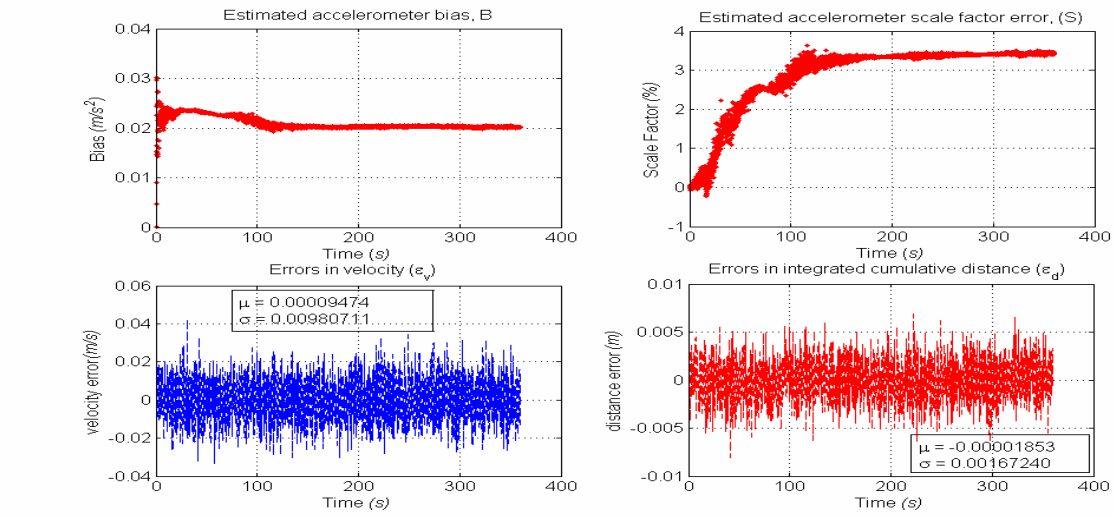

Figure 6.19: *Effects of the simulated wheel speed sensors' and accelerometer noise (Test 1A)*

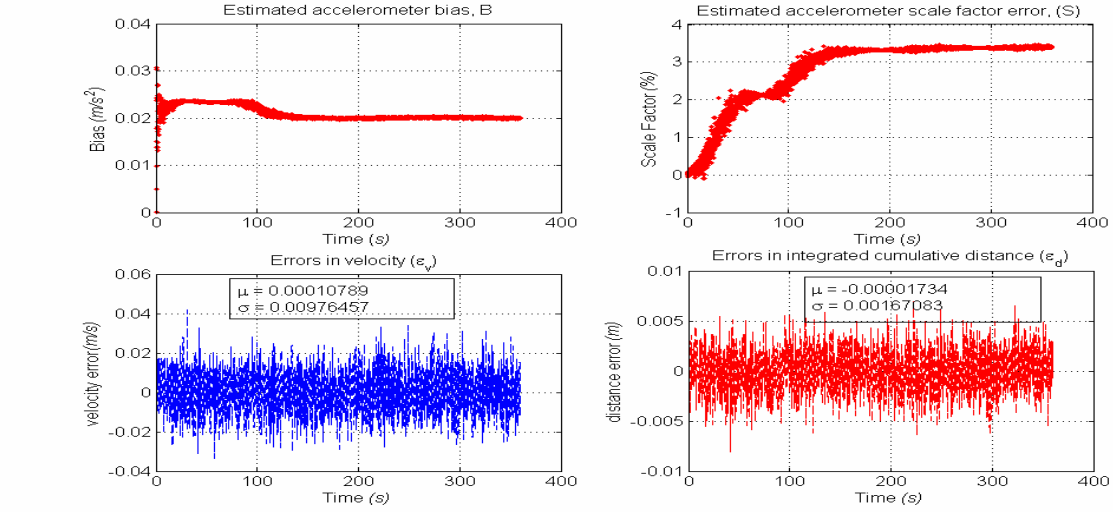

Figure 6.20: *Effects of the simulated wheel speed sensors' and accelerometer noise (Test 1B)*

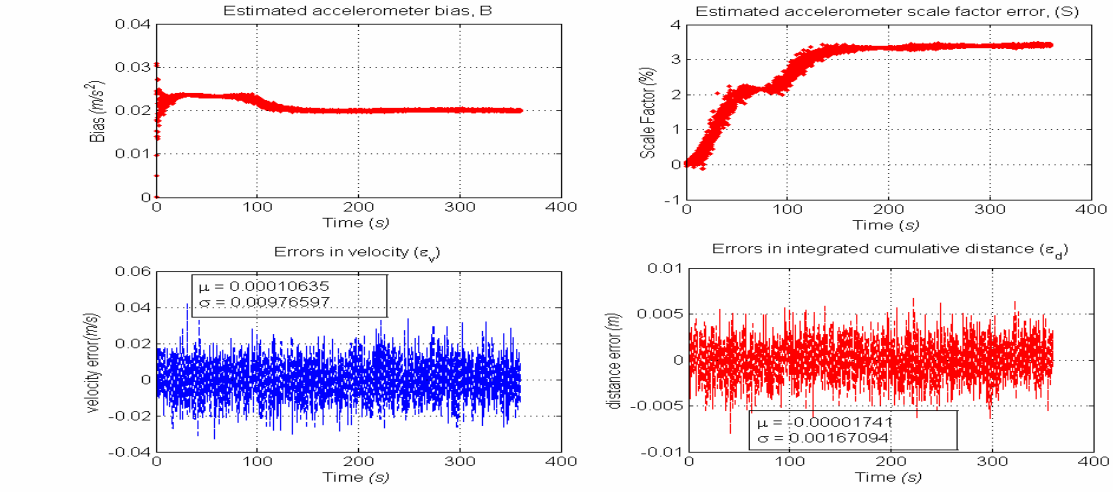

Figure 6.21: *Effects of the simulated wheel speed sensors' and accelerometer noise (Test 1C)*

Table 6.9: *Parameter settings used in testing the effects of simulated ABS measurement noise*

	Test 2A	*Test 2B*	*Test 2C*
Accelerometer Bias/ scale error	0.020 m/s^2 / 3.5%	0.020 m/s^2 / 3.5%	0.020 m/s^2 / 3.5%
Accelerometer Measurement Noise (m/s^2)	$N \sim (0, 0.000625)$	$N \sim (0, 0.000625)$	$N \sim (0, 0.000625)$
Initial bias-covariance	0.64	0.64	0.64
ABS Measurement Noise (m) (m/s)	Distance: $N \sim (0, 0)$ Velocity: $N \sim (0, 0)$	Distance: $N \sim (0, 0.000185^2)$ Velocity: $N \sim (0, 0.00185^2)$	Distance: $N \sim (0, 0.00185^2)$ Velocity: $N \sim (0, 0.0185^2)$
Initial scale factor-covariance	0.64	0.64	0.64

For the second set of tests, the standard deviation of the random noise in accelerometer measurements ($\sigma = \pm 0.025 m/s^2$) is held constant while the standard deviations of the simulated normally distribution random noise in the wheel speed sensor measurement are set to $\sigma = \pm 0.0m$ in *Test 2A*, $\sigma = \pm 0.000185m$ in *Test 2B* and $\sigma = \pm 0.00185m$ in *Test 2C* as summarised in Table (6.9). Similarly using an initial state vector variance of *0.8 m/s* in speed and *8.8 m* in distance, the estimated accelerometer biases and scale factor errors are shown in Figures (6.22), (6.23) and (6.24). From the results, the standard deviation of the estimated velocities increases with the increase in wheel speed sensor noise from $\sigma = \pm 0.00054 m/s$ in *Test 2A,* $\sigma = \pm 0.00057 m/s$ in *Test 2B,* to $\sigma = \pm 0.00188 m/s$ in *Test 2C.* Similarly the standard deviation of the estimated distances increases with the increase in wheel speed sensor noise from $\sigma = \pm 0.000063m$ in *Test 2A,* $\sigma = \pm 0.000066m$ in *Test 2B,* to $\sigma = \pm 0.000212m$ in *Test 2C.* By comparing the two sets of results, one sees that the influence of wheel sensors random noise is relatively higher than the effect of accelerometer random noise. If the characteristics of the random noise are known, the information may be incorporated into the Kalman filter model. The challenge therefore is to characterization of the system random errors.

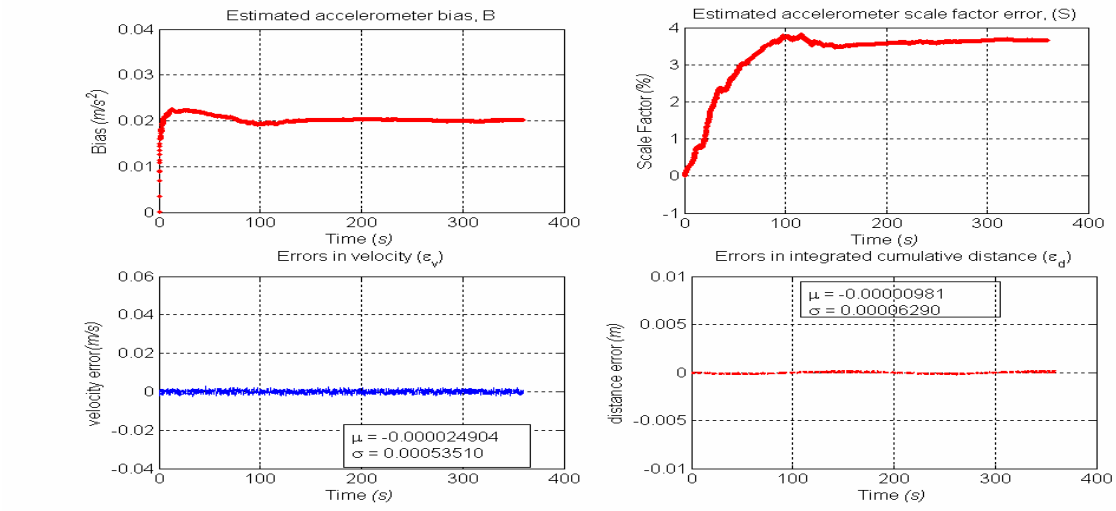

Figure 6.22: *Effects of simulated accelerometer noise (Test 2A)*

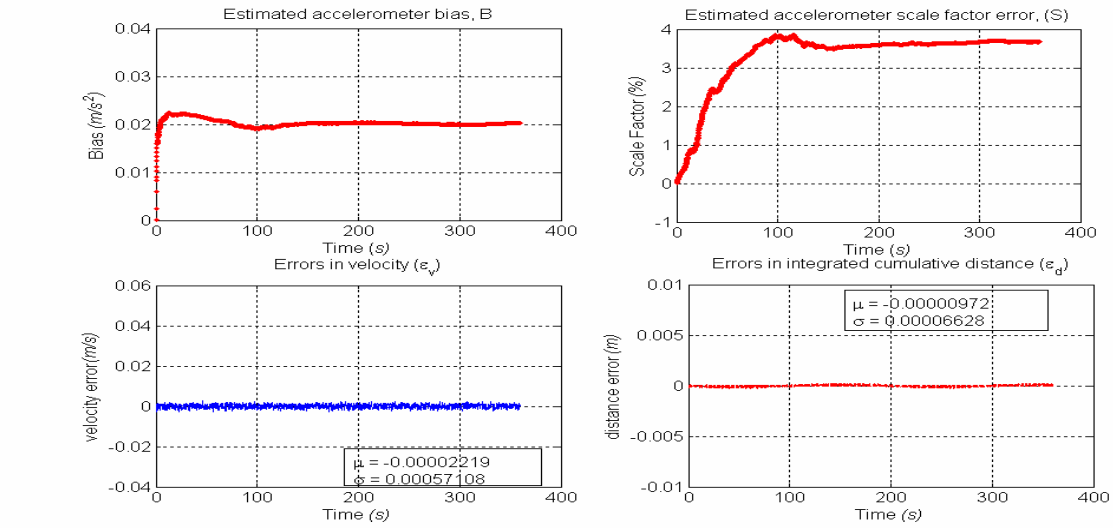

Figure 6.23: *Effects of the simulated wheel speed sensors' and accelerometer noise (Test 2B)*

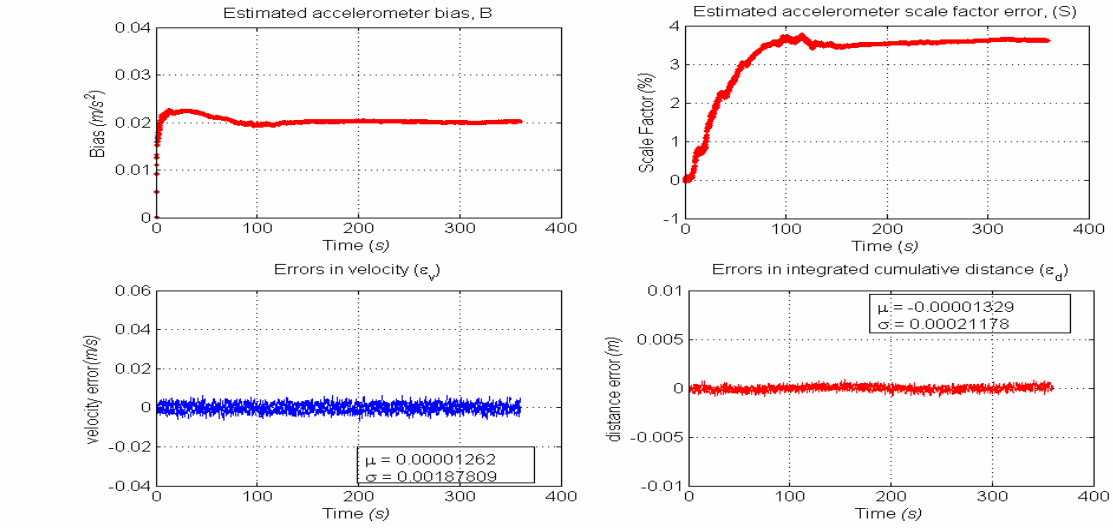

Figure 6.24: *Effects of the simulated wheel speed sensors' and accelerometer noise (Test 2C)*

Maintaining the same settings for the bias error at *0.02 m/s²* , the scale factor error at *3.5%* , the variance of the simulated random noise in accelerometer measurements and ABS measurements as in the case of first set of tests summarized in Table (6.8), the integration of the ABS wheel speed sensor and accelerometer measurements was performed by invoking the threshold speed which is set at *0.367 m/s* as defined in Section (6.2.1.3). The results of the estimated accelerometer biases and scale factor errors are shown in Figures (6.25) through (6.27). By invoking the threshold speed condition, the errors in velocity and distances increase tremendously with respect to time. For the three tests, the estimated maximum deviation in speed and distance errors reached maximum values of *0.2 m/s* and *0.9 m* respectively during the measurement outages. As expected, the Kalman filter effectively reduces the effects of accelerometer noise on distance and velocities and therefore in the absence of the update measurements (filtering), the accelerometer data cannot be used.

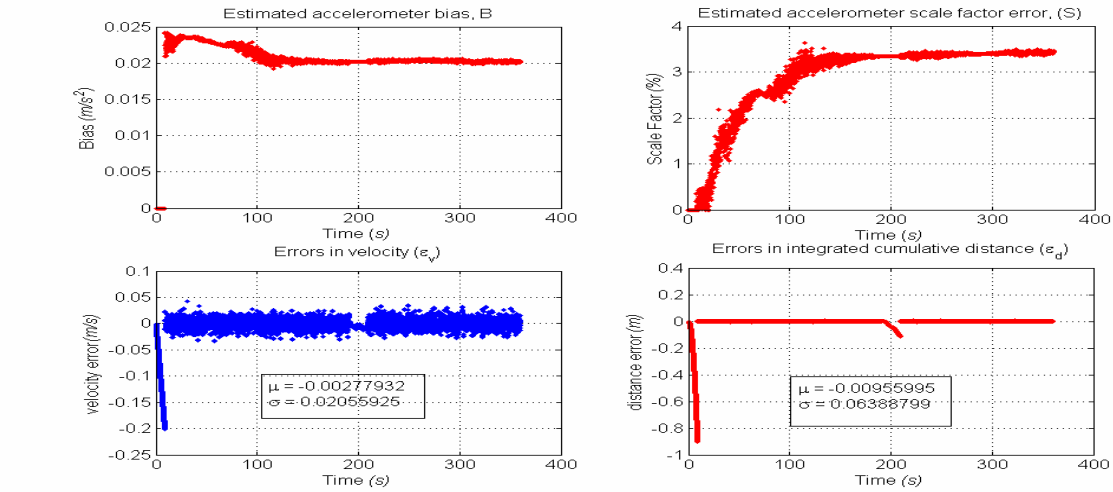

Figure 6.25: *Effects of simulated ABS and accelerometer measurement noise with threshold speed condition (Test 1A)*

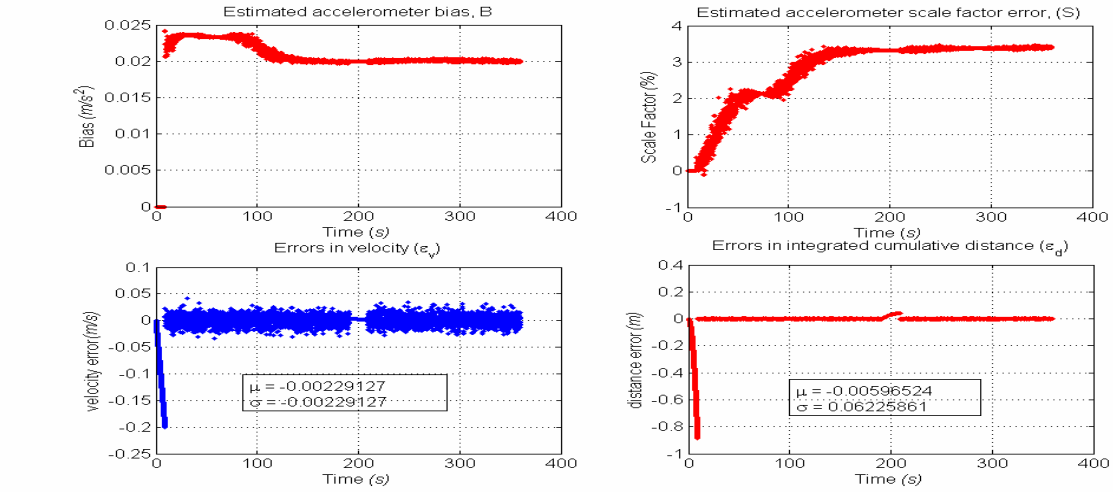

Figure 6.26: *Effects of simulated ABS and accelerometer measurement noise with threshold speed condition (Test 1B)*

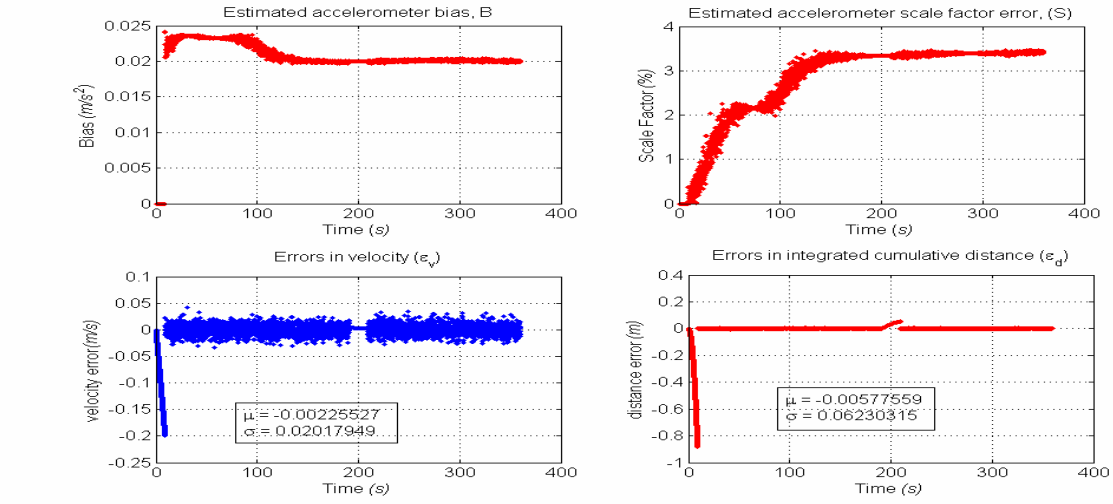

Figure 6.27: *Effects of simulated ABS measurement noise with threshold speed condition (Test 1C)*

Again, maintaining the same settings for the bias error at *0.02 m/s²* , the scale factor error at *3.5%* and the variance of the simulated random noise in accelerometer measurements and ABS measurements as in the case of first set of tests summarized in Table (6.9), the integration of the ABS wheel speed sensor and accelerometer measurements is performed by invoking the threshold speed which was set at *0.367 m/s* as defined in Section (6.2.1.3). The results of the estimated accelerometer biases and scale factor errors are shown in Figures (6.28) through (6.30). By invoking the threshold speed condition, the errors in velocity and distances increase tremendously with respect to time. For the three tests, the estimated maximum deviation in speed and distance errors reached maximum values of *0.2 m/s* and *0.9 m* respectively during the measurement outages. And again as expected, the Kalman filter effectively reduces the effects of accelerometer noise on distance and velocities and therefore in the absence of the update measurements (filtering), the accelerometer data cannot be used.

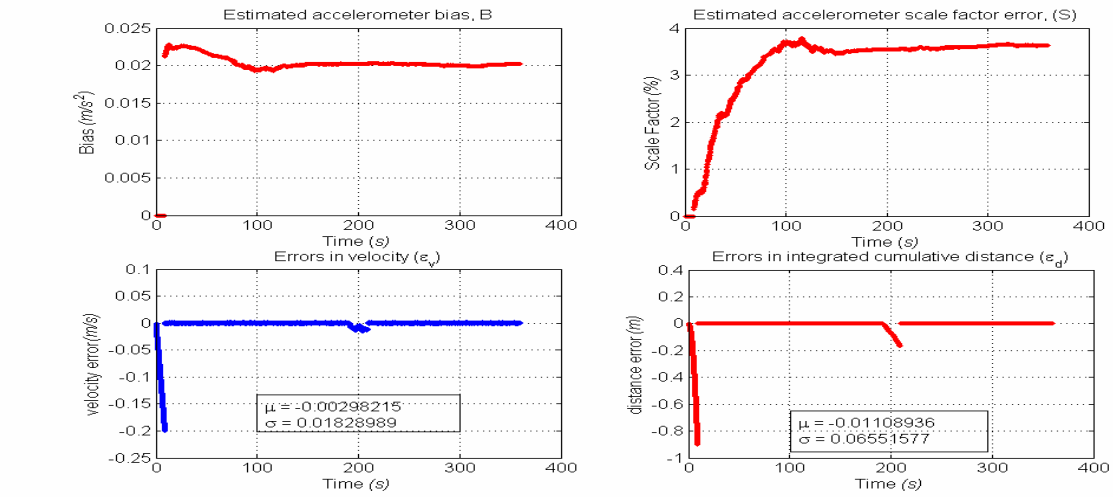

Figure 6.28: *Effects of simulated accelerometer measurement noise with threshold speed condition (Test 2A)*

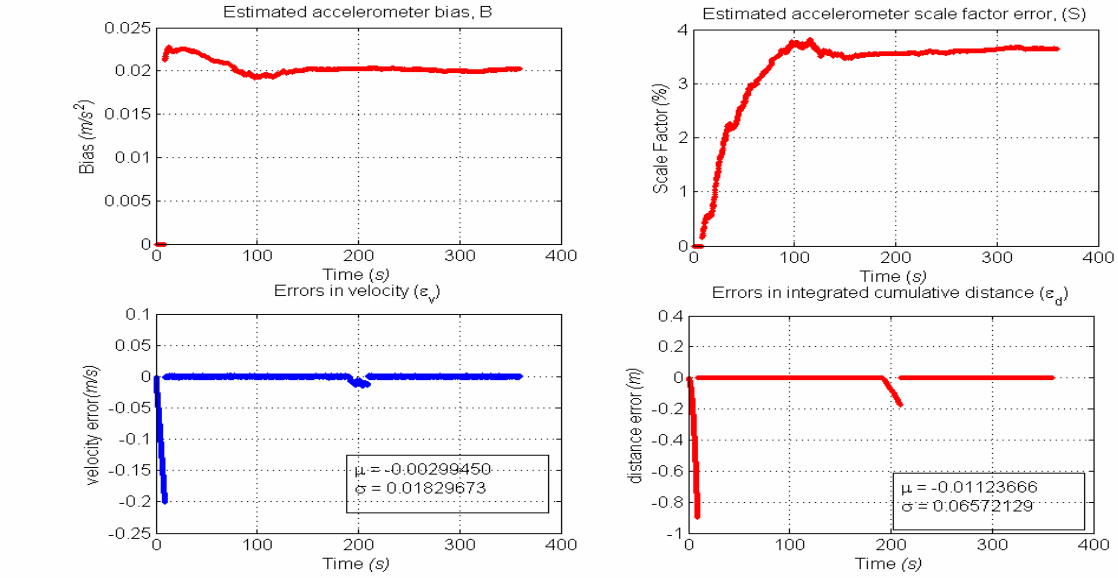

Figure 6.29: *Effects of simulated ABS and accelerometer measurement noise with threshold speed condition (Test 2B)*

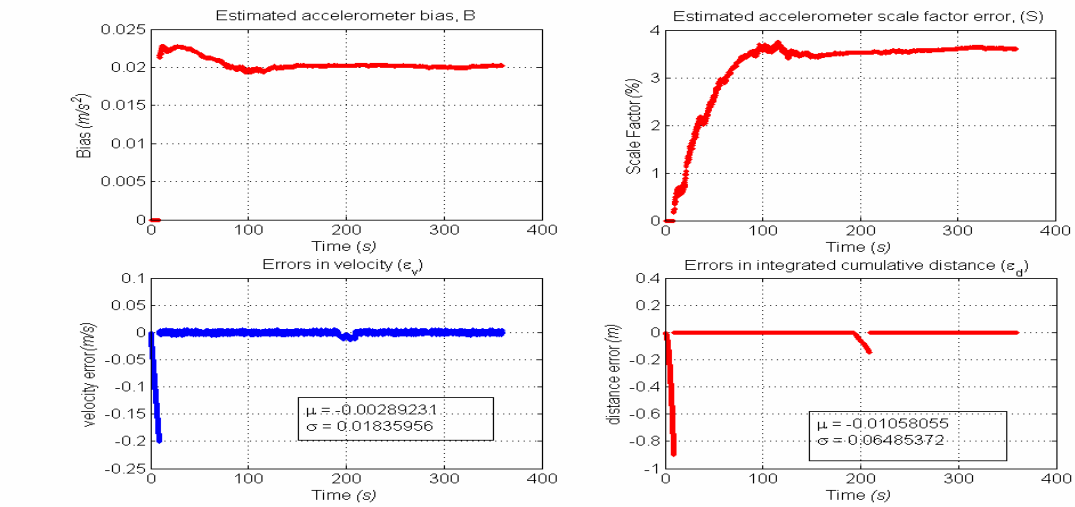

Figure 6.30: *Effects of simulated ABS and accelerometer measurement noise with threshold speed condition (Test 2C)*

6.3.3 Real Data

Figures (6.31) and (6.32) show the *20 minute* long *100 Hz* measured cumulative distance and speed of the vehicle during the test drive in the vicinity of Stuttgart city. The Applanix *POS proc* software solutions of the Applanix POS AV™ measurements are used as true values to test and analyse the lost-cost sensors' measurements. Figures (6.33), on the other hand, shows the *100 Hz* raw *x*-axis acceleration measured by the MotionPak™ measured accelerations.

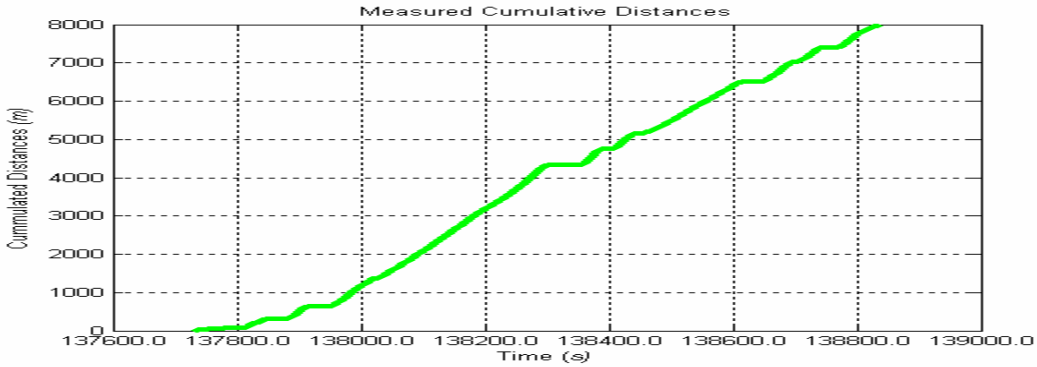

Figure 6.31: *Applanix POS AV measured cumulative distances of the test drive*

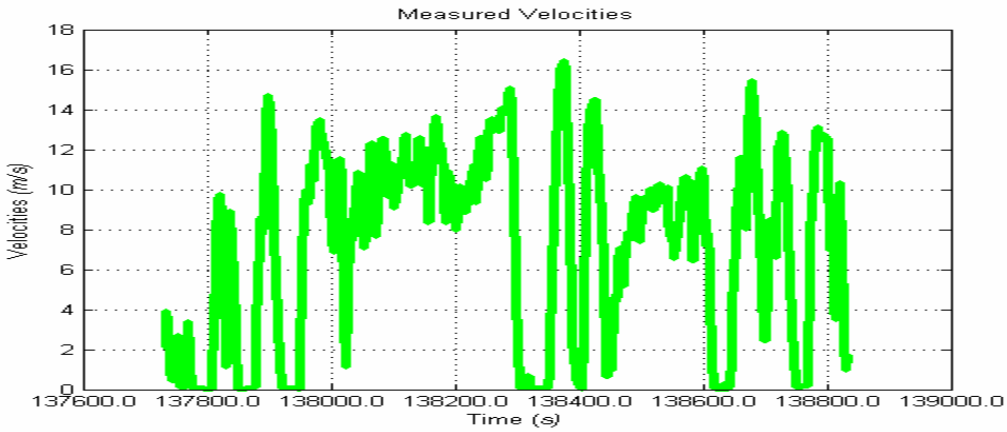

Figure 6.32: *Applanix POS AV™ measured velocities of the test drive*

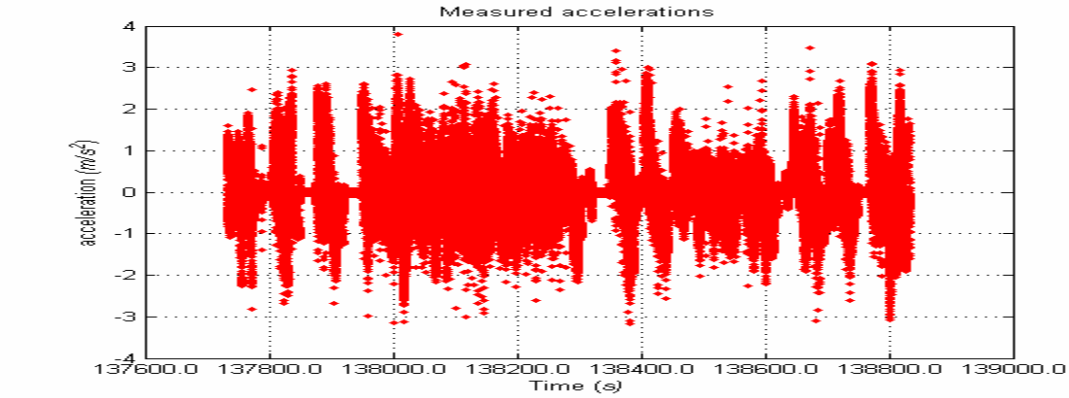

Figure 6.33: *MotionPak™ measured acceleration values of the test drive*

6.3.3.1 Real data integration results

In order to estimate the noise level in the accelerations measured using the MotionPak™, an auto regression moving average analysis is performed. Table (6.10) shows the result of the analysis.

Table 6.10: *Auto regression moving average (ARMA) results (m/s²)*

order	Real data (stationery vehicle)		Real data (entire data set)	
	Corrected for Pitch & gravitation	noisy	Corrected for Pitch & gravitation	noisy
56	0.000087034	0.00026056	0.0722318	0.077197
255	0.000073141	0.00023909	0.0718246	0.076761
1255	0.000067128	0.00022757	0.0706707	0.075509

Tuning of the filter

Measurement noise covariance assumed after auto regression moving average (ARMA) analysis of the ABS measurement noise is given by;

$$CV_n = \begin{bmatrix} 0.0185^2 & 0 \\ 0 & 0.185^2 \end{bmatrix} \tag{6.61}$$

The initial estimated covariance of the state vector is taken as;

$$\Sigma_{A_0} = \begin{bmatrix} 0.8^2 & 0 & 0 & 0 \\ 0 & 0.8^2 & 0 & 0 \\ 0 & 0 & 0.8^2 & 0 \\ 0 & 0 & 0 & 0.8^2 \end{bmatrix} \tag{6.62}$$

Process noise covariance was assumed as in Equation (6.63).

$$\Sigma_B = \begin{bmatrix} 0.05^2 & 0 & 0 \\ 0 & 0.00005^2 & 0 \\ 0 & 0 & 0.00005^2 \end{bmatrix} \tag{6.63}$$

Table 6.11: *Parameter for estimation of accelerometer bias and scale factor error*

	Test 3A	*Test 3B*	*Test 3C*
Accelerometer Measurement Noise	$N \sim (0, 0.00025)$	$N \sim (0, 0.00025)$	$N \sim (0, 0.00025)$
Initial bias-covariance	0.64	0.64	0.64
ABS Measurement Noise	Distance: $N \sim (0, 0.0185^2)$ Velocity: $N \sim (0, 0.185^2)$	Distance: $N \sim (0, 0.00185^2)$ Velocity: $N \sim (0, 0.0185^2)$	Distance: $N \sim (0, 0)$ Velocity: $N \sim (0, 0)$
Initial scale factor-covariance	0.64	0.64	0.64

The ABS measurements are used in three tests and are assumed to be contaminated with normally distributed random noise with a mean of zero and a variance of 0.0185^2 in distances and 0.185^2 in velocities while the accelerations were assumed to be contaminated with normally distributed random noise with a mean of zero and a variance of 0.0025^2 in the first filter (*Test 3A*), a variance of 0.00025^2 in the second filter (*Test 3B*) and a variance of 0.000025^2 in the third filter (*Test 3C*) as summarised in Table (6.11). The results in the estimated accelerometer biases and scale factor errors are shown in Figures (6.34) (6.35) and (6.36). One sees from the analysis that the reduction in measurement covariance does not influence the estimation of the Bias and scale factor error significantly.

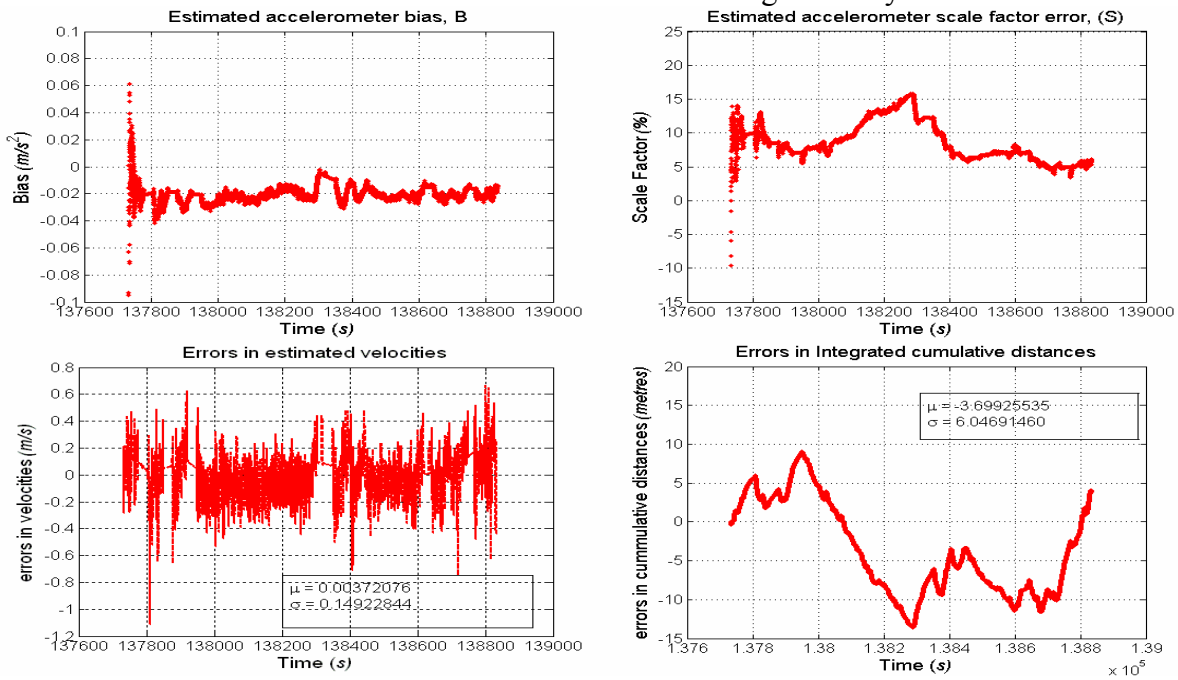

Figure 6.34: *Effects of accelerometer and ABS measurement noise (Filter 3A)*

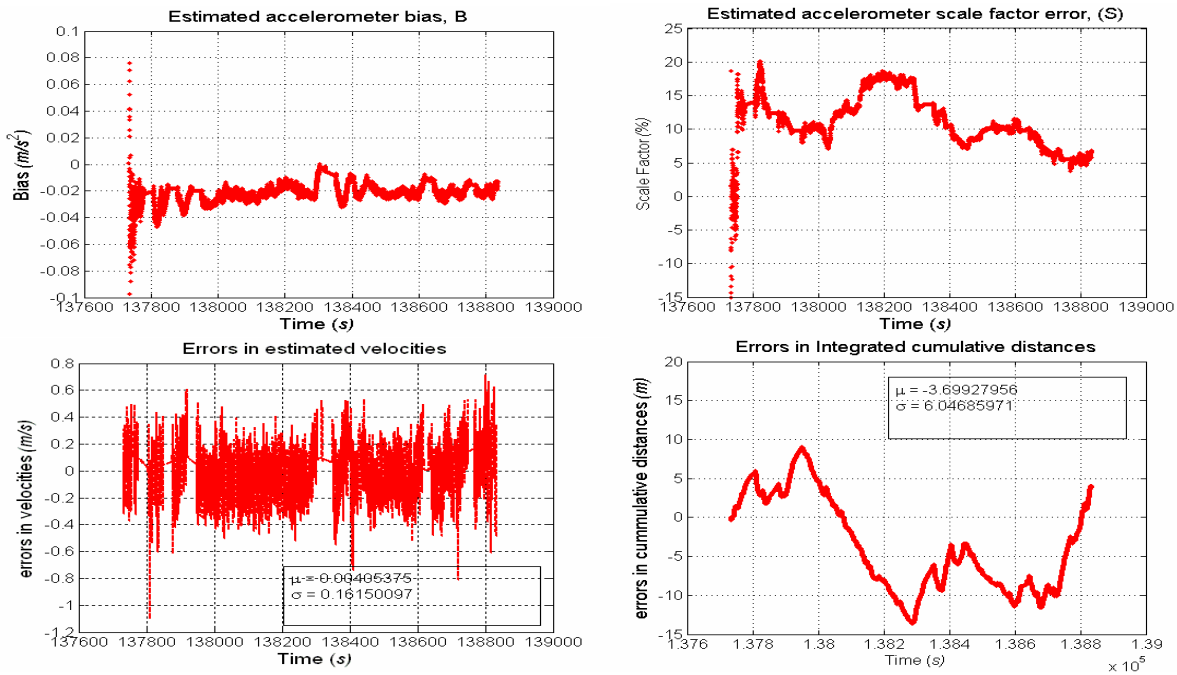

Figure 6.35: *Effects of accelerometer and ABS noise (Filter 3B)*

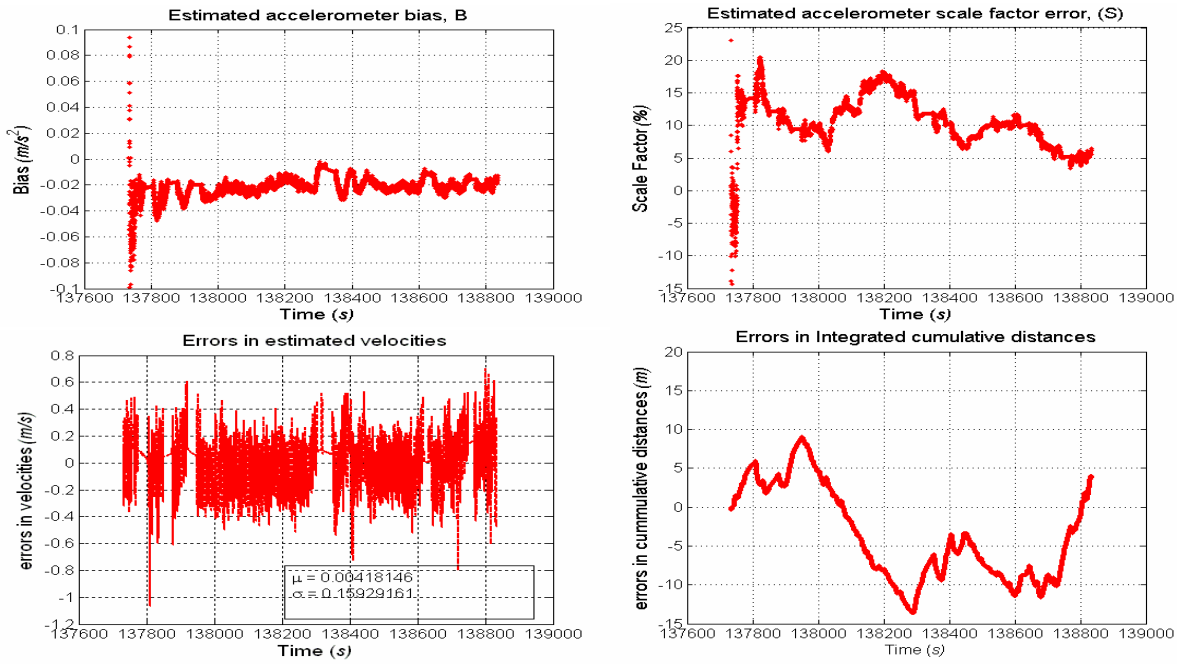

Figure 6.36: *Effects of accelerometer and ABS noise (Filter 3C)*

From the figures, one concludes that the wheel speed sensors have an accuracy of about *0.04 m/s* in velocity determination and maximum deviation of about *10 m* in the cumulative distance derived from the measurements. Performing the same integration but invoking the threshold speed to 0.*367 m/s* results in the estimated accelerometer biases and scale factor errors as shown in Figures (6.37) (6.38) and (6.39). As can be seen from the figures, the random noise in the ABS derived distances and velocities have varied effects on the estimated bias and scale factor error as opposed to the case with the simulated data. The effects of threshold speed are similar in behaviour.

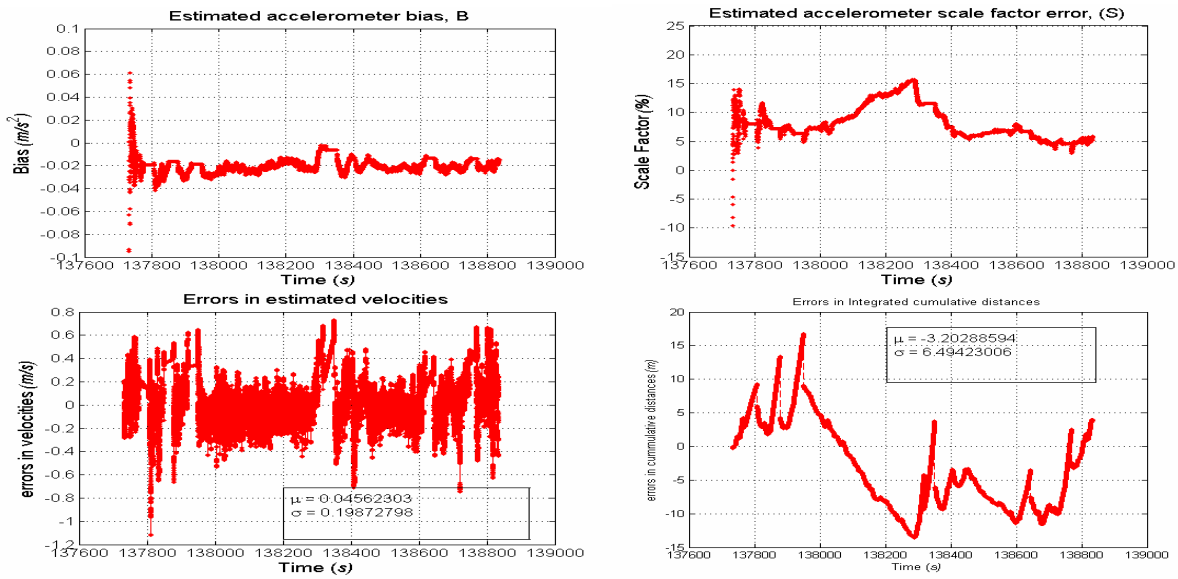

Figure 6.37: *Effects of accelerometer and ABS noise with threshold speed (Filter 3A)*

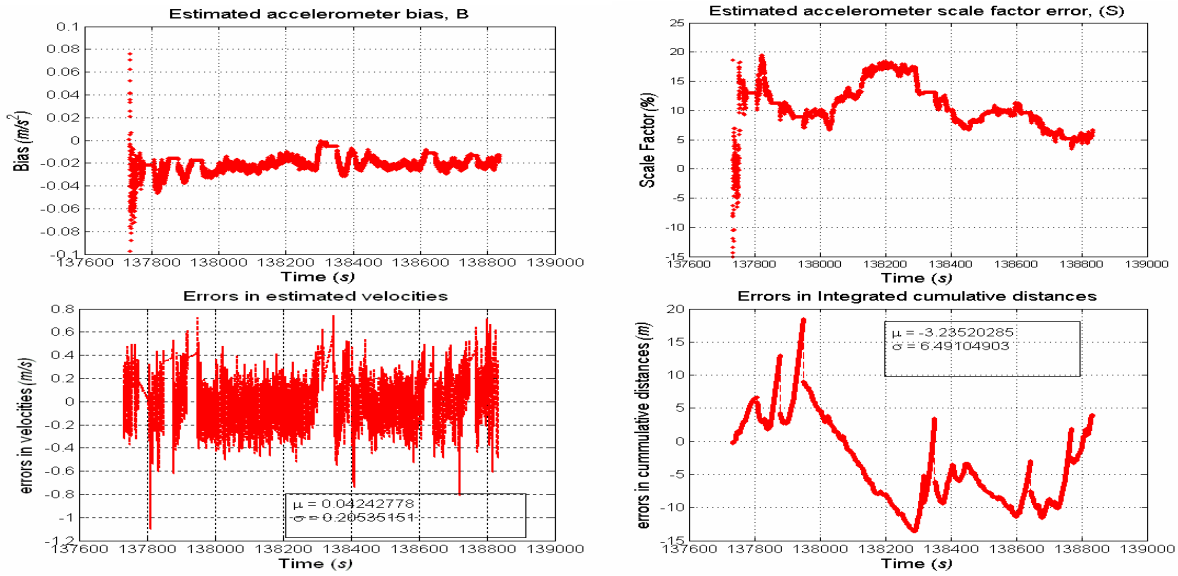

Figure 6.38: *Effects of accelerometer and ABS noise with threshold speed (Filter 3B)*

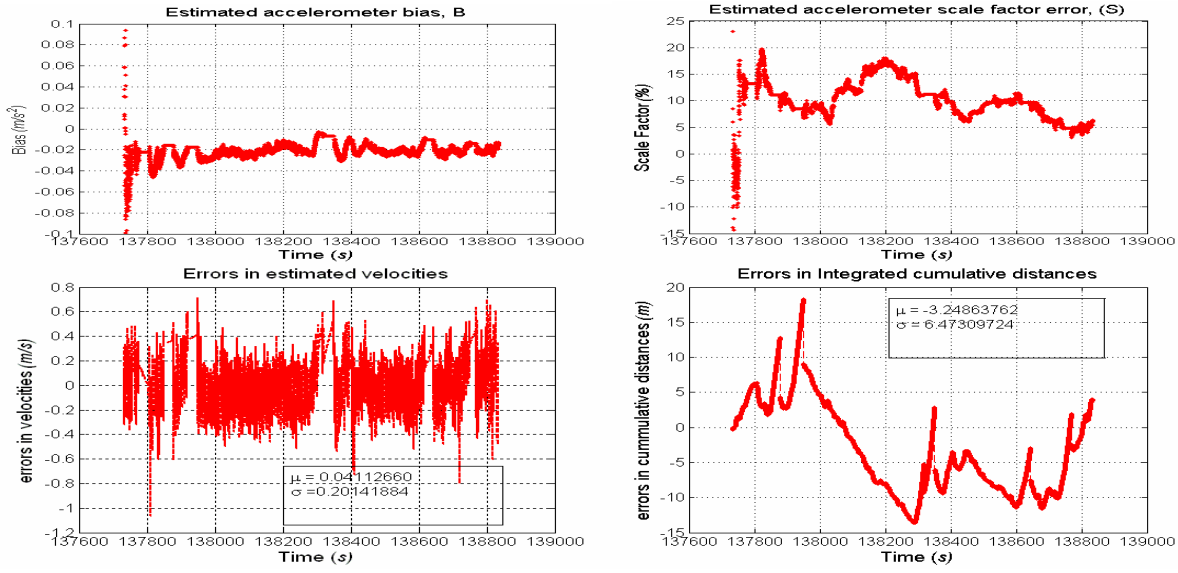

Figure 6.39: *Effects of accelerometer and ABS noise with threshold speed (Filter 3C)*

6.3.4 Final Remarks on integration of ABS and accelerometer measurements

One may conclude from the analysis of the data that

- the fluctuations in the estimated measurement updates in the Kalman filtering process are dependent on the setting of the process covariance. Setting higher process covariance values allows for higher fluctuations and vice versa.
- measurements are assumed to have constant covariance which may not be true in the real case situation due to several other localized noise variations.
- initial state vector covariance settings have little effect on the final estimated parameters and their covariance but will determine the period required by the filter to be optimal. As the measurement error covariance approaches zero, the Kalman gain increases hence the measurements are heavily weighed.
- attempts to use accelerometer data to bridge ABS measurement gaps during the period when the velocity is below the threshold speed is not successful due to large accelerometer errors which accumulate in the absence of the ABS measurements. Thus the ABS measurements used so long as the velocity is maintained above the threshold speed, the acceleration is maintained at relatively constant values and the accuracies of in distance and velocities are assumed to be or better than $\pm 4cm$.
- The Kalman filter based algorithm has been proved to achieve the set goal as expected and the accuracy of the derived distances and velocities are of acceptable accuracy for such a low cost system.

6.4 Estimation of the vehicle heading and curvature

Integration of the raw *z*-axis angular velocity measured by the MotionPak™ gyro shows very large drift of up to $+0.25°/s$ as compared to true vehicle heading measured by the Applanix POS AV system. Two main corrections are applied to the measured angular rates. First, the bias drift is estimated from averaged gyro measurements while the vehicle is stationary then applied to the subsequent measurements. Secondly, pitch angle corrections are applied using the pitch angles derived from the Applanix system. The MotionPak derived heading error was then tremendously reduced to $\pm 3°$ as shown in Figures (6.40) and (6.41). The pitch angle is therefore an important parameter in heading estimation using single axis gyros. Application of single-axis gyro in heading estimation requires a pitch angle sensor such as an inclinometer for the purposes of pitch angle and gravitational corrections.

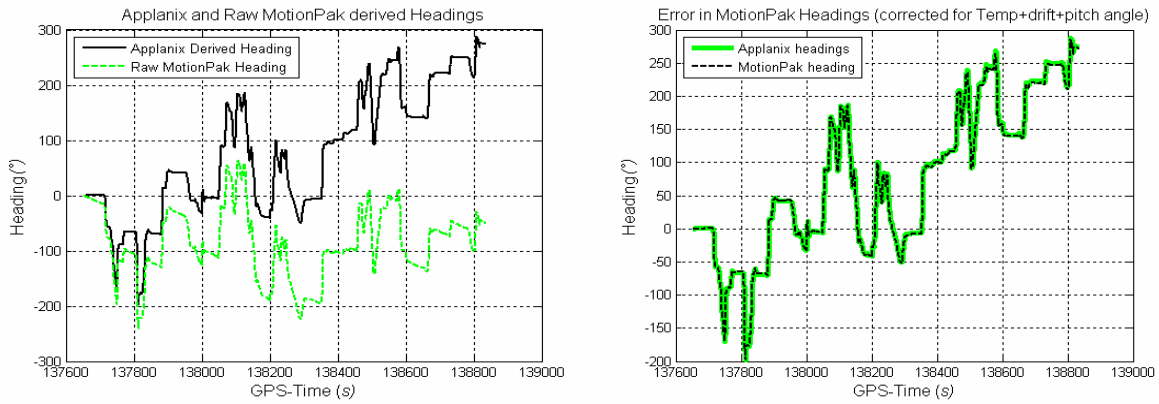

Figure 6.40: *Raw and corrected heading (Applanix and gyro derived)*

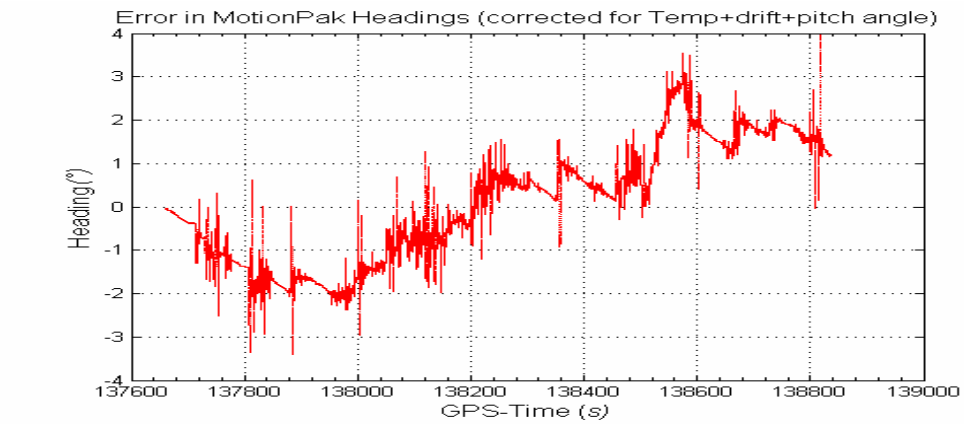

Figure 6.41: *Error in corrected gyro derived heading*

The estimated vehicle heading together with the distance increments derived from the ABS and accelerometer integration are used in the computation of the vehicle positions. To determine whether the errors in the estimated headings or the distance increments results in significant error in position along the track, three computational analyses are carried out. The true vehicle heading estimated by the Applanix system is used with the estimated distance increments to estimate the vehicle position and position errors as shown in Figure (6.42). The maximum position error is about *9 m* in both *x*- and *y*-directions. This translates to about *0.14 %* position error for the *7 kilometre* track length.

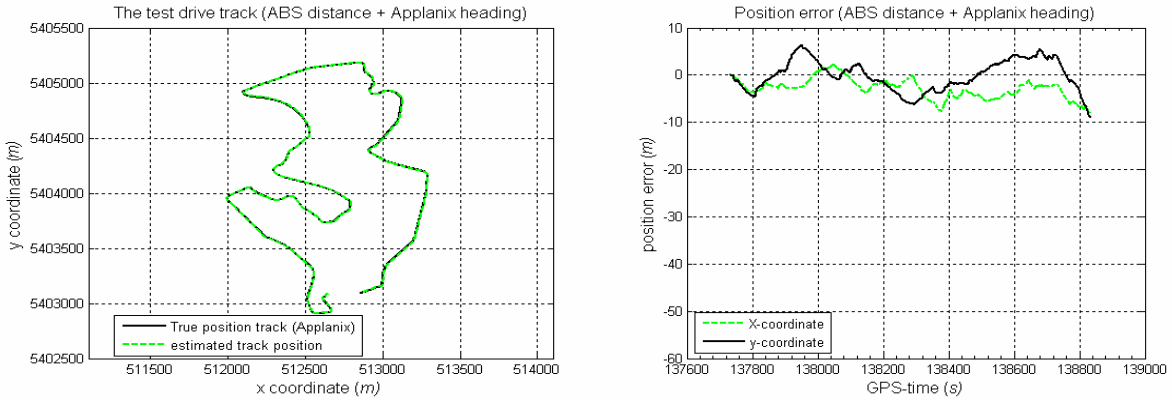

Figure 6.42: *Derived position and position errors (ABS distances and Applanix heading)*

The true distance increments estimated by the Applanix system are again used together with the MotionPak derived heading differences to derive the vehicle position along the track. The position and position errors obtained are shown in Figure (6.43) with a maximum position error of *42 m* in the *x*- direction and *15 m* in the *y*-direction. In comparison to previous results, one clearly sees that the heading results in the most significant position error along the track.

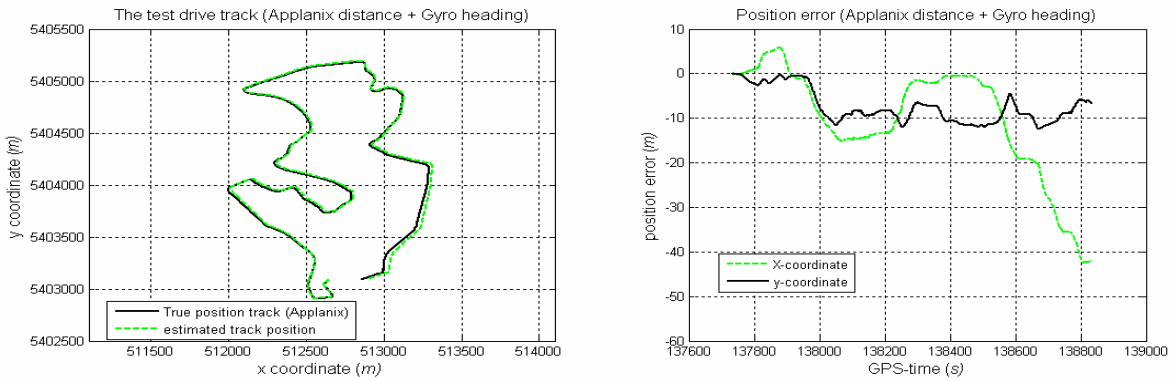

Figure 6.43: *Derived position and position errors (gyro heading and Applanix distance)*

By using both the estimated distance increments and heading differences i.e. estimation of the track position independent of the Applanix derived distances and heading. The position and position errors obtained are shown in Figure (6.44). The resulting maximum position error is about *55 m* in the *x*-direction and *15 m* in the y-direction. Again, the significant contributions to position error are the estimated heading differences.

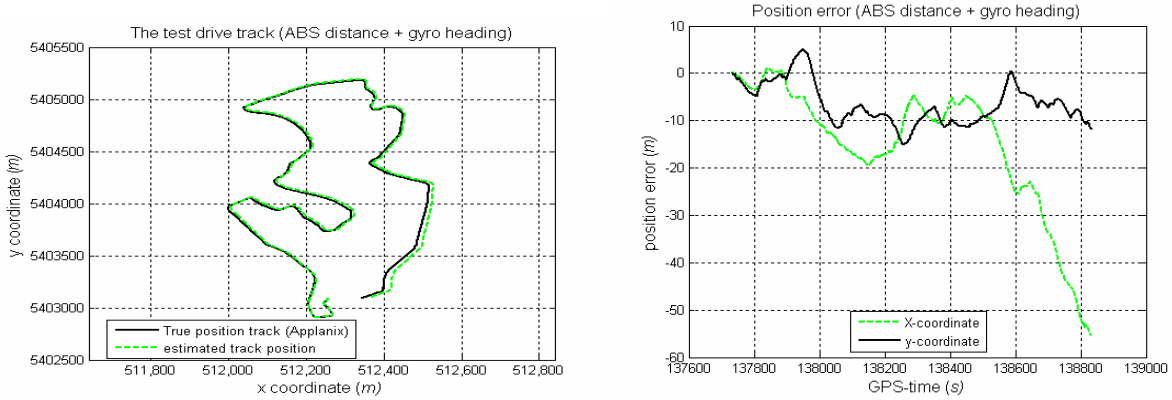

Figure 6.44: *Derived position and position errors (gyro heading and ABS distance)*

The curvature values computed from the ABS and accelerometer derived distance increment and single-axis gyro derived heading differences are shown in Figure (6.45) while the errors in the estimated curvatures in comparison to Applanix POS AV derived curvature are shown in Figure (6.46). The resulting standard deviation of about $\pm 0.0000025 m^{-1}$ translates to errors in the order of *5 to 10 m* in the recovered radius of curvatures of between *500 to 1000 m* respectively or in other words, results in *1%* error of the radius of curvature.

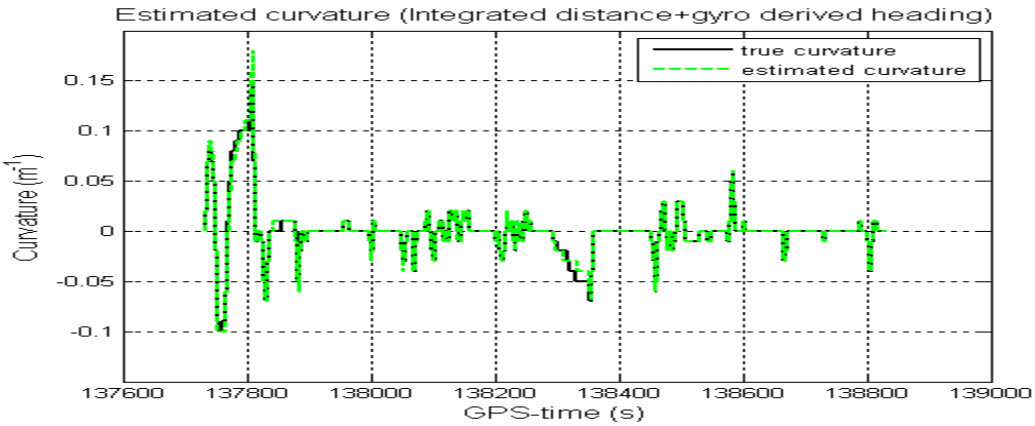

Figure 6.45: *Estimated curvature values*

Figure 6.46: *Error in estimated curvature values*

From the analysis, the proposed low-cost systems results have large position errors but relatively small errors in curvature allowing the recovery of the geometric shape of track. These navigation sensors used in deriving the geometry can be combined with a low cost GPS receiver to allow the estimation of the positions of the recovered geometric design elements.

7 Summary and Conclusions

Data collection process for the purposes of developing a roadway inventory database is an expensive undertaking by various road agencies around the world, whose main responsibility is the initial development, storage and management of the database. The roadway inventory database comprises of several categories of collectable data elements which include for example horizontal curve elements (radius and curve length), the vertical alignment elements (tangent grades and lengths), drainage, number of lanes and traffic signs along the road among others. As a result of the rapidly changing roadway environment, the challenge therefore is the establishment of relatively accurate and low-cost data collection and processing methods and devices for updating the different roadway inventory database elements. The work presented in this thesis illustrates the low-cost sensors layout, the methodologies used in the field measurement and the algorithms developed for the extraction of the geometric road design parameters of the vehicle track for the above purpose. The algorithms are developed and tested in the MATLAB programming language environment and designed to carry out the following tasks; synchronisation of the measurements and calibration of the measuring sensors, estimation and reduction of the deterministic and stochastic errors in the vehicle track measurement, sensors integration and vehicle track curvature estimation respectively recovery of the geometric road design elements from the position time series. The user developed algorithms are located in the hard disc with the flowcharts showing some of the key components of the algorithms listed and described in the Appendices.

In summary, Chapter (5) demonstrate that the presence of the deterministic and stochastic errors in the measured vehicle track position time series, are a serious predicament to curvature estimation as proved by the analysis of their combined effects as illustrated in Section (5.5). Although the influences of the systematic errors on the estimated curvature value are relatively higher than the effects of random errors, the former can be eliminated through the calibration processes of the measuring sensors and systems. The application of cubic smoothing splines has been shown to reduce the influence of the random errors on curvature estimation but the same failed to have significant effects on the estimation of the positions of the starting and ending points of the recovered geometric parameters. The smoothing factor used by the cubic smoothing splines in the filtering process was determined from the standard deviation of the measurements, obtained by characterisation of the random errors in the field measurements. Meanwhile, Chapter (6) presents the algorithms and methodologies developed for the sensors calibration, sensors integration, curvature estimation and recovery of the geometric road design elements as tested using computer simulated and field measured data. The field calibration procedure of the ABS wheel speed sensors achieved an accuracy of $\pm 0.04m$ and since the accuracies of the geometric road design elements in the roadway inventory database are specified in sub-meter, the accuracy of the distance increments derived from the ABS wheel speed sensors, obtained after synchronisation and with the GPS data, was considered to be sufficient for the stated research objectives.

Low-cost integrated navigation system

A set of low-cost sensors was applied in the measurement of the vehicle track positions. A linear state system modelled for the ABS and accelerometer integration through Kalman filtering process was developed and used to estimate the accelerometer scale factor and bias errors. The results of the integrated ABS and accelerometer derived distance increments

demonstrate that use of the accelerometer measurements to bridge the ABS wheel speed sensors signal errors (signal truncations), when the vehicle speed falls below the threshold speed, do not meet the performance requirement due to rapid accumulation of the errors. From the analysis of the distance increments derived from the integrated ABS and accelerometer system set-up shown in Figure (6.32), the accuracy of the accumulated distance is reduced to about $\pm 6m$ during the signal truncations. Failure of the accelerometer measurements to bridge the ABS wheel sensors data gaps may also be due to highly vibrating road environment. By analysing the field data collected using the measuring van driven around Stuttgart city for about *20* minutes and covering a track length of about *7* kilometres, the derived distance increments and single-axis derived heading demonstrate that the system can be used for measuring position time series of the vehicle track for the recovery of the geometric road design parameters. A deviation of about $\pm 13\ m$ in cumulated vehicle track length of *7* kilometres translates to error in distance of about *0.2%* or *2000 ppm* while the derived heading difference have a final cumulative deviation of $\pm 3°$ or a drift of about *0.0025°/s* after the pitch angle correction and gravitation compensation in the gyro measurements. In conclusion, the study has demonstrated that such a set of low-cost sensors consisting of ABS wheel speed sensors, a single-axis MEMS accelerometer and a single-axis MEMS gyro is a viable solution as a short duration georeferencing device for mobile mapping systems in the recovery of the geometric shape of the road track, so long as (i) the vehicle is driven at low speed (about *36 km/h or 10 m/s*) but certainly above the threshold speed during the measurements, (ii) the characterisation of the random errors in the measuring sensors and (iii) the compensation for gravitation is performed on the measured data during post processing.

Curvature estimation and recovery of geometric road design elements

Relatively accurate position time series of the vehicle track (road) are required in order to derive the track geometry. The position time series used in the estimation of road geometry of other linear feature of the roadway inventory database are obtained either *directly* by use of land-based integrated navigation systems or *indirectly* through automatic recognition and point following of the georeferenced scanned images usually by airborne systems. Two approaches are then normally used in the estimation of the geometry of the road from the estimated position time series. One approach uses the entire position time series of the measured trajectory while the second approach uses a predictive concept in which the data are analysed sequence by sequence to derive the required parameters. The advantage of the first approach is than it allows for assessment and correction to the vehicle movement thus resulting in smoothly adjusted parameters and is performed in a post processing procedure. Although the latter approach starts off with fewer data and possible ill-suited due to fluctuations, filtering could be applied to reduce the fluctuations. The second approach can be suitable for the real-time intelligent transport systems.

The curvatures estimated using the developed algorithms were analysed and tested successfully using the simulated data as can be seen in Figure (5.3). Unfortunately, clearly defined geometric sequence of straight line, circular and transitional curves was not achieved using the field data measured around Stuttgart city centre. Analysis of the curvatures estimated from the field data, as presented in Figures (6.45) and (6.46) illustrate the feasibility of the set of low-cost sensors as a tool of determining the geometric shape of the vehicle track to a sufficient accuracy despite the presence of deterministic and stochastic errors in the field measured position time series. In conclusion, the results demonstrate that the set of the low-cost sensors together with the developed spline-based algorithms successfully extracted the

geometric road design elements for updating of the roadway inventory database based on the assumption that the road geometry consist of straight lines, circular and transitional curves. The geometric road design elements may though not be sufficiently accurate for road design and construction purposes but can support the intelligent transport system applications such as driver assistance and mobile information services.

Based on the analysis of the computer simulated data and field measurements, one recommendation is that the improvements of the technique can be achieved by integrating the ABS derived distance increments and gyro derived heading with low-cost GPS receiver data. The field measurements were carried in the vicinity of Stuttgart city where the roads are designed without observing the geometric road design requirements. It would be interesting to observe the results of the same tests on rural roads or motorway. Another recommendation is a retrial of the experiment using optical gyro which does have mass and is therefore not affected by the dynamics of the environment in which they are being applied in this case the roadway environment. It is also worth noting that the pitch angle correction plays a greater role in the processing of the single-axis gyro measurements and therefore the incorporation of an inclinometer into the sensors systems is required. In the application of the cubic splines in the curve fitting and curvature estimation processes, the curves were parameterised using curve length, a modification of the algorithms so that the curve is parameterized by time i.e. generating unequal distance intervals due to varying vehicle speed should be investigated and analysed. Last but not least, development of the algorithms into computer programs interfaced with the graphical user interface (GUIs) is planned.

References:

American Association of State Highway and Transport officials (AASHTO)(1994): A policy on Design of Highways and Streets. http://www.aashto.org/.

Arbeitsgruppe Straßenentwurf (ed.), 1996: *Richtlinie für die Anlage von Straßen (RAS), Teil Querschnitte (RAS-Q 96)*. Forschungsgesellschaft für Straßen und Verkehrswesen, Bonn.

Babu, R. and J. Wang (2005): *Ultra-tight GPS/INS/PL integration: Kalman filter performance analysis*. GNSS 2005, Hong Kong, 8-10 December.

Barbour, N. And G. Schmidt (2001): Inertial Sensor Technology Trends. IEEE Sensors Journal, Vol. 1, No. 4, December.

Bernstein, J (2003): *An overview of MEMS inertial sensing technology*. Sensors digital edition, http://www.sensorsmag.com/articles/0203/14/main.shtml.

Bonnifat, Ph. and G. Garcia (1998): *Design and Experimental validation of an Odometric and Goniometric Localisation System for outdoor Robot Vehicles*. IEEE Transactions on Robotics and Automation, Vol. 14, No. 4, August'98, pp. 541-548.

Bonsen, G., D. Ammann, M. Ammann, E. Favey and P. Flammant (2005): *Continuous Navigation: Combining GPS with Sensor-Based Dead Reckoning*. GPS world, pp, April.

Borenstein, J. (1998): *Experimental Evaluation of a Fibre Optics Gyroscope for Improving Dead-reckoning Accuracy in Mobile Robots*. In proc. of the IEEE Int. Conference on Robotics and Automation, Leuven, Belgium, May 16-21, pp. 3456-3461.

Brandmeier, T. (2004): *Simulative Zuverlässigkeitsbewertung für Sicherheitskritische Fahrzeugsysteme*. ASIM-Symposium in FH-Inglostadt, 9-10 March, 2004. http://www-rt.upb.de/ASIM/FG455/Inglostadt/FG_Manuskripte/brandmeier.pdf

Brockenborough, R. L. and K. J. Boedecker (1996): *Highway Engineering Handbook: Building and Rehabilitating the Infrastructure*. ISBN 0-07-008777-6, McGraw Hill, NY, USA.

Carlson, C. R., J. C. Gerdes and J. D. Powell (2003): *Practical position and Yaw Rate Estimation with GPS and Differential Wheelspeeds*. Proc. of the AVEC, February.

Caspary, W., (2002): *Kinematische Messmethoden*. Zfv 127. jg. 5/2002.

Caspary, W., H. Heister, C. Hock, J. Klemm and H. Sternberg (1995): *A Kinematic Survey System for Transportation Infrastructure*. In Proc. FIG-Symposium Committee VI, 36-41, Cape Town.

CEN TC 287, (1995): *Geographic data files version 3.0*, GDF for Road Traffic and Transport Telematics, Technical report, CEN.

Centre for Mapping Website, Ohio State University, http://www.cfm.ohio-state.edu/research/itscurve.html

Christiani Technical Institute for Vocational Training (2005), http://www.kfztech.de/kfztechnik/fahrwerk/bremsen/abs.htm

Chung, H., L. Ojeda and J. Borenstein (2001): *Accurate Mobile Robot Dead- Reckoning with a Precision-Calibrated Fibre-Optic Gyroscope*. IEEE Transactions on Robotics and Automation, Vol. 17, No. 1, February, pp. 80-84.

Coconis, P. (2005): *ABS Speed Sensor Service, A clean sensor is a happy sensor*. Retrieved on 25th October 2005 Web: http://www.advanceautoparts.com/english/youcan/html/

Coetsee, J., A. Brown and J. Bossler (1994): *GIS Data Collection using GPS Van supported by a GPS/Inertial Mapping System*. In proc. of ION GPS-94, Salt Lake City, UT, 20-23 September.

Cox, M. G. (1972): *The numerical evaluation of B-splines*. Jour. Inst. Math. Applic., Vol. 10, pp. 134-149.

Cox, M. G. and Hayes, J. G. (1973): *Curve fitting: a guide and suite of algorithms for non-specialist user*. NPL Report NAC 26, National Physical Laboratory.

Curry, H. B. and Schoenberg, I: J. (1947): *On spline distributions and their limits: the Pólya distribution functions,* Abstract 380t, Bull. Amer. Math. Soc., Vol. 53, p. 109.

De Boor, C. (1972): *On calculating with B-splines*, Jour. Approx. Theory, Vol. 6, pp. 50-62.

De Boor, C. (1978): *A practical guide to Splines*, Springer-Verlag, New York.

Dierckx, P. (1975): *An algorithm for smoothing, differentiating and integration of experimental data using spline functions*. Jour. Comput. Appl. Math. Vol. 1, pp. 165-184.

Dierckx, P. (1981): *An improved algorithm for curve fitting with spline functions*. Report TW54, Department of Computer Science, Katholieke Universiteit Leuven.

Dierckx, P. (1982): *A fast algorithm for smoothing data on a rectangular grid while using spline functions*. SIAM J. Numer. Anal. 19, pp. 1286-1304.

Dierckx, P. (1995): *Curve and surface fitting with splines*. Clarendon Press, Oxford.

Dorota, A., Grejner-Brzezinska and C. K. Toth (2002): *Multisensor System for Automatic Monitoring of Highway Linear Features.* FIG XXII International Congress, Washington D. C. U.S.A., April.

El-Sheimy, N. (2001): *Summary report on the proceeding.* The 3rd International Symposium on Mobile Mapping Technology, Cairo, Egypt, January 3-5.

El-Sheimy, N. (2003): ENGO 623 Lecture Notes: *Inertial Techniques and INS/GPS Integration*. Department of Geomatics Engineering, University of Calgary, Winter 2003.

Farin, G. (1997): *Curves and Surfaces for Computer Aided Geometric Design*. 4th ed., Academic Press, San Diego, California.

Farrell, J.A and M. Borth (1999): *The Global Positioning System and Inertial Navigation*. McGraw-Hill, New York.

Forder, R. (2000): *Design and development of a flying platform*. MEng Thesis,

Forschungsgesellschaft für Straßen- und Verkehrswesen e.V. (1995): *Richtlinien für die Anlage von Straßen (RAS)*: Teil *Linienführung (RAS-L)*, Köln.

Forschungsgesellschaft für Straßen- und Verkehrswesen (FSV) e.V. (1988): *Richtlinien für die Anlage von Straßen (RAS)*: Teil *Leitfaden für die funtionale Gliederung des Straßennetzes* (RAS-N), Köln.

Forschungsgesellschaft für Straßen- und Verkehrswesen (FSV) e.V.(1985): *Empfehlungen für die Anlage von Erschließungsstraßen (EAE)*; Köln.

Forschungsgesellschaft für Straßen- und Verkehrswesen (FSV) e.V.(1993): *Empfehlungen für die Anlage von Hauptverkehrsstraßen (EAHV 93)*, Köln.

Forschungsgesellschaft für Straßen- und Verkehrswesen (FSV) e.V.(1982): *Richtlinien für die Anlage von Straßen (RAS)*, Teil: *Querschnitte (RAS-Q)*, Köln.

Forschungsgesellschaft für Straßen- und Verkehrswesen (FSV) e.V.(1992): Hinweis zur Anwendung der *Richtlinien für die Anlage von Straßen*, Teile: *(RAS-N, -L, -Q, -K) beim Um- und Ausbau von Straßen in der neuen Bundesländern*, Köln.

Gao, J. (2002): *Dynamic position sensing for parallel kinematic machine*. PhD Thesis, The University of Nottingham.

Gao, J., P. Webb and N. Gindy (2003): *Error reduction for inertial-sensor-based dynamic parallel kinematic machine positioning system*. Meas. Sci. Technol. 14, (2003), Institute of Physics publishing, pp. 543-550

Gat, E. (1995): *Towards Principled Experimental Study of Autonomous Mobile Robots*. ISERS95

Gil, J., and D. Keren (1997): *New approach to the arc length parameterization problem*. 13th Spring Conference on Computer Graphics, pp. 27-34.

Gray, A. (1997): *Curvature of Curves in the plane, in Modern Differential Geometry of Curves and Surfaces with Mathematica*. 2nd ed. Boca Raton, FL:CRC Press, pp 14-17.
Greiner, G. (2002): *Lecture on Geometric Modelling* (with scripts and maple worksheets), http://www9.informatik.uni-erlangen.de/Teaching/WS2002/GM
Guerrero, H., J. L. Escudero and E. Bernabeu (1994): *A magneto-optical tacheometer for anti-lock systems using plastic optical fibre*. Meas. Sci. Technol. 5 (1994) 607-610. UK, January 1994.
Hartley, P. J. and C. J. Judd (1980): *Parameterization and Shape of B-Spline Curves for CAD*. CAD, 12(5), September, p.235
Hatger, C. and C. Brenner (2003): *Extraction of road geometry parameters from laser scanning and existing databases*. Worksop on 3-D Reconstruction from Airborne Laser Scanner and InSAR Data, Dresden, Germany, October 8-10.
Hay, C. (2005): *Turn, Turn, Turn Wheel Speed Dead Reckoning for Vehicle Naviagtion*. GPS World, pp 37-42October 2005.
Hayes J. G. (1974): *Numerical Methods for curve and Surface Fitting*. Bull. Inst. Math. Appl. 10, 144-152.
Hinueber, E. V. (2006): If you are investigating Inertial Measurement Systems. iMAR Website: http://imar-navigation.de/
Höllig, K. (2001): *Finite Element Methods with B-Splines*. Society for industrial and applied mathematics (SIAM), Philadelphia, U.S.A, ISBN 0-89871-533-4.
Hou, H. (2004): Modeling *Inertial Sensors Errors Using Allan Variance*. Msc. Thesis, Department of Geomatics Engineering, University of Calgary, UCGE Reports Number 20201.
Housedolder, A. S. (1953): *Principles of Numerical analysis*. McGraw Hill Book Co., New York, pp 196-200.
Ingenieurgesellschaft für kinematische Vermessung (IKV) Website, http://www.ikv-kiss.de/index.htm
Jekeli, C. (2001): *Inertial navigation systems with geodetic applications*. Walter de Gruyter, Berlin, New York.
Jöns, O. (2001): *Genauigkeits- und Stabilittsanalyse der MotionPak™ –Innertialmesseinheit*. Studentarbeit II, Institut für Navigation, University of Stuttgart.
Joy, K. I. (1999), *Definition of a B-Spline Curve. On-line Geometric Modelling notes*. Department of computer science, UCD, California, USA.
Kalman, R. E. (1960): *A new approach to linear filtering and prediction problems*. Transaction of the AMSE-Journal of basic Engineering, pp. 35-45, March
Kelly, A. (1996): *Position Estimation 3: Sensors for Position Estimation*. Fall.
Khattak, A. J., J. E. hummer and H. A. Karimi (2001): *Evaluation of Roadway Infrastructure Data Collection Technologies*. Public Works Management & Policy, Vol. 6, N. 1, pp. 18-31.
Kleusberg, A. (2004): *Introduction to Navigation I and II*. Unpublished Lecture Notes, Institute of Navigation, University of Stuttgart.
Kluge, K. and C. E. Thorpe (1999): *Representation and Recovery of Road Geometry in YARF*. In Proc. Symposium on Intelligent Vehicles '92, Detroit, USA, June 29 – July 1, pp. 114-119.
Kobryn, A. (1991): *Zur Berechnung von Absteckdaten bei der Anwendung der Blosscurve*. Zeitschrift für Vermessungswesen, Germany, 91/7,pp. 284-288.
Landstar (2005): website http://www.landstar.com/
Lightsey, G. (2006): Lightweight Low Cost Integrated GPS/INS Navigation. Lightsey research group website http://gps.csr.utexas.edu/GPS-INS/
Mckinley, S. and Levine, M. (1995): *Cubic Spline Interpolation*. Math 45: Linear Algebra, http://online.redwoods.cc.ca.us/instruct/danold/laproj/fall9/skymeg/proj.pdf

Meng, X. (1999): *Vehicle-borne Highway Geometric Alignments and Facilities Data Capture Using DGPS and GPS/GIS Integration*. In: Proc of ION GPS'99 (student paper prize), Nashville, 14-17 September, 1999.

Merhav, S. (1996): *Aerospace Sensor Systems and Applications*. Springer-Verlag, New-York, Inc.

Mölzer, P. and R. Zurek (1977): *Trassierungselemente für Schnellbahnen mit elokromagnetischer Schwebetechnik*. ETR4.

Mostafa, M. M. R., J. Hutton, B. Reid, R. Hill (2001): *GPS/IMU Products* – the Applanix Approach, In Fritsch/spiller (eds.): Photogrammetric Week 2001, Wichmann Verlag, Heidelberg, Germany, p. 63-83.

Muskett, J. (1995): *Site Surveying*. 2nd ed, Blackwell Science, Oxford, UK.

Nayak, R. A. (2000): *Reliable and Continuous Urban navigation Using Multiple GPS Antennas and a Low Cost IMU*. Master thesis, Department of Geomatics Engineering, University of Calgary, UCGE Reports No. 20144, October.

Niu, X., C. Goodall, S. Nassar and N. El-Sheimy (2006): *An Efficient Method for Evaluating the Performance of MEMS IMUs*. IEEE/ION PLANS, San Diego, California, USA, April 24-27, pp. 766-771.

Nowak, K.(1991): *Integration von GPS und digitalen Kameras zur automatischen Vermessung von Verkehrswegen. Zeitisch*. F. Photogr. U. FernerK., S. 112-120.

Oberlechner, G., B. Metzger, J. Zywiel (2000): *GPS System To Boost Accuracy*, International Railway Journal, September.

Ogonda, G. O (2003): *Setting up of a reference station and investigation of the effect of a random*. Master of science Thesis, Institute of Navigation, University of Stuttgart

Ogonda, G. O (2005): *Calibration of the ABS wheel speed sensors*. Unpublished internal report, Institute of Navigation, University of Stuttgart.

Ojeda, L., H. Chung and J. Borenstein (2000): *Precision-calibration of Fibre-optics Gyroscopes for Mobile Robot Navigation*. Proc. of the 2000 IEEE Int. Conf. on Robotics and Automation, San Francisco, CA, April 24-28, pp. 2064-2069.

Omnistar (2005): website http://www.omnistar.com/

Pappas, I. P. I., T. Keller and M. R. Popovic (1999): *Experimental Evaluation of Gyroscope Sensor used in a New Gait Phase Detection System*. IFESS99, 4th Annual Conference of the International Functional Electrical Stimulation Society, August 23-27, Sendai, Japan.

Peterson, J. W. (2006): *Arc Length Parameterization of Spline Curves*. www.saccade.com/writing/graphics/RE-PARAM.pdf

Piegl, L. and Wayne, T. (1997); *The NURBS book*, 2nd edition, Springer Verlag, Berlin, Germany, ISBN 3-540-61545-8.

Psarianos, B., D. Paradisis, B. Nakos and G. Karras (2001): *A Cost-effective Road Surveying Method for the Assessment of Road Alignments*. Proceedings: IV International Symposium Turkish-German Joint Geodetic Days, Berlin, April 3-6.

Ramshaw, L. (1987): *Blossoming: A connect-the-dots approach to splines*, Report 19, Digital Systems Research Centre, Palo Alto, CA.

Ressel, W. (1999): *Einfuhrung in die Vekehrsplanung* Teil: Straßenwesen, Vorlesungsumdruck WS 1999/2000, Lehrstuhl Straßenplanung und Straßenbau, ISV, Universität Stuttgart.

Ribbens, W. (1994): *Understanding Automotive Electronics*. SAMS Publishing

Salychev, O. S., V. V. Voronov, M. E. Cannon, R. Nayak and G. Lachapelle (2000): *Low-Cost INS/GPS Integration: Concepts and Testing*. ION National Technical Meeting, Anaheim, CA, Jan 26-28.

Sharpe, R. J. and R. W. Thorne (1982): *Numerical Method for Extracting an Arc Length Parameterization from Parametric Curves.* CAD, 14(2), March, p.79.

Shene C. K. (2006): *Introduction to Computing with Geometry.* CS3621 Notes, Department of Computer Science, Michigan Technological University. http://www.cs.mtu.edu/~shene/COURSES/cs3621/NOTES/notes.html

Scherzinger, B.M. (1993): *A Report on Navigation Systems Integration and on the Benefits and Costs of Integration*. Unpublished report of the Applied Analytics Corporation, Markham, Ontario, Canada. Prepared for the Directorate of Avionics, Systems and Photography, Department of National Defense of Canada.

Schoenberg, I. J. (1946a): *Contributions to the problem of approximation of equidistant data by analytic functions, Part A: on the problem of smoothing or graduation*, *a first class of analytic approximation formulae*, Quart. Appl. Math., Vol. 4, pp. 45-99.

Schoenberg, I. J. (1946b): *Contributions to the problem of approximation of equidistant data by analytic functions, Part B: on the problem of oscillatory interpolation*, *a second class of analytic approximation formulae*, Quart. Appl. Math., Vol. 4, pp. 112-141.

Schwarz, K. P., N. El-Scheimy, R. Li, and M. A. Chapman (1994): VISAT – *A Mobile System for Precise GIS Highway Inventory Applications*. In Proc. 3rd Int. Conf. on Land Vehicle Navigation, 283-297, Dresden, Germany.

Schwarz, K. P. and N. El-Sheimy (1999): Future Positioning and Navigation Technologies. Study performed under the Scientific Services Agreement with Batelle, Columbus Division and Topographic Engineering Center, Fort Belvoir, VA, USA.

Sennott, J. and D. Senffner (1997): *Robustness of Tightly Coupled Integrations for real-Time Centimeter GPS Positioning*. 10th Int. Tech. Meeting of the Satellite Division of the U.S. Inst. of Navigation, Kansas City, Missouri, 16-19 September, pp. 655-663.

Sennott, J. (1999): *Receiver Architectures for Improved Carrier Phase Tracking in Attenuation, Blockage and Interference*. GPS Solutions, 3, pp. 40-47.

Skaloud, J.,A. M. Bruton and K. P. Schwarz (1999): Detection and Filtering of Sort Term (1/f) Noise in Inertial Sensors. Journal of Institute of Navigation, Vol. 46, No. 2, pp 7-107.

Smith, R. B. and J. R. Weyrauch (1990): *Gyroscopes: Current and Emerging Technologies*. IAG Symposium on Kinematic Systems in Geodesy, Surveying and Remote Sensing, No. 107, Alberta, Canada, September 10-13.

Stathas, D., J. Karabelas, C. Liapakis, B. Psarianos and M. Kontaratos (2001): *Dynamic Vehicle Positioning and Orientation Techniques for Inspecting Road Surface or Guiding Construction Machines*. New Technology for a New Century International Conference, FIG Working Week 2001, Seoul, Korea, May 6-11.

Sternberg, H. and W. Caspary (2002): *Fahrzeugtrajektorie und Trasse als Ergebnisse kinematischer Vermessung.* In Zeitschrift für Vermessung (Jg. 127).

Sternberg, H., W. Caspary, H. Heister and J. Klemm (2001): *Mobile Data Capturing on Roads and Railways Utilizing the Kinematic Survey System KiSS.* The 3rd International Symposium on Mobile Mapping Technology, Cairo, Egypt, January 3-5.

Systron Donner (2006): Inc. website (http://www.systron.com/pro_MotPk.asp)

TRB Circular, 1999: *Research Directions For In-Vehicle Computing: Delivering an Integrated Suite of ITS Applications for Driver Assistance and Mobile Services & Information*. http://gulliver.trb.org/publications/circulars/ec009.pdf, Transport Research Board, Washington D.C., No. E-C009, June.

Titterton, D. H. and J. L. Weston (1997): *Strapdown Inertial Navigation Technology*. In IEE Radar, Sonar, Navigation and Avionics Series 5, E Bradsell, Ed. London, U.K., Peters Pergrinus Ltd., pp 455-455.

Titterton, D. H. and J. L. Weston (2005): *Strapdown Inertial Navigation Technology*. American institute of Aeronautics and Astronautics, Reston, VA, USA.

Tribbett et al. (2000): *An Evaluation of Dynamic Curve Warning Systems in the Sacramento River Canyon*: Final Report of the Western Transport Institute, Montana State University, http://www.coe.montana.edu/wti/wwwshare/curve_warning/FINAL REPORT.pdf ,April.

US DoD (2001): *Global Positioning System Standard positioning Service Performance Standard*. A Report of the U.S Assistant Secretary of Defence, October.

Vaidya, A., D. Baraniak and K. Novak (1994): *CONRAILS's GIS Incorporates Kinematic GPS and Stereo Digital Imaging*. In AM/FM Inernational Conference Proceedings, Aurora, CO.

Van Zeghbroeck, B. (1998): *Hall effect and Summary*. ECEN5355 Lecture#6, 4th September.

Wang, H., J. Kearney and K. Atkinson (2002): *Arc-Length Parameterized Spline Curves for Real-Time Simulation*. Presented at the 5th International Conference on Graphics and Image Processing, pp. 113-135.

Waters N. M. (1999): *Transportation GIS: GIS-T*. In Longley P, Goodchild M, Maguire D, and Rhind D (eds) Geographical Information Systems: Management Issues and Applications. John Wiley & Sons, Inc, USA.

Weisstein, E. W. (2006a): *Bézier Curve*. From Mathworld - A Wolfram Web Resource, http://mathworld.wolfram.com/BezierCurve.html ,April.

Weisstein, E. W. (2006b): *B-Spline*. From Mathworld - A Wolfram Web Resource, http://mathworld.wolfram.com/B-Spline.html ,April.

Weisstein, E. W. (2006c): *Curvature*. From Mathworld - A Wolfram Web Resource, http://mathworld.wolfram.com/Curvature.html ,April.

Weisstein, E. W. (2006d): *Conic Section*. From Mathworld - A Wolfram Web Resource, http://mathworld.wolfram.com/ConicSection.html ,April.

Whittaker, W. and L: Nastro (2006): *Utilization of Position and Orientation Data for Preplanning and Real Time Autonomous Vehicle Naviagtion*. Position Location and Navigation Symposium (PLANS 2006), IEEE/ION, CA, April 25-27.

Wikipedia (2006): The free encyclopaedia http://en.wikipedia.org/wiki/spline_interpolation

Wright, P. H., (1996): *Highway Engineering*. 6th ed., ISBN 0-471-00315-8, John Wiley & Sons, NY, USA.

Young, S. E. (2004): *Development of a Continuous Three Dimensional Geometric Highway Model using Historical Global Positioning System Data*. Doctoral dissertation, Kansas State University.

Zhao, L., W. Y. Ochieng, M. A. Quddus and R. B. Noland (2003): *An Extended Kalman Filter Algorithm for Integrating GPS and Low Cost Dead Reckoning System Data for Vehicle Performance and Emissions Monitoring*. The Journal of Navigation (2003), 56, pp. 257-275, U.K.

Zhao, Y. (1997): *Vehicle location and navigation systems*. ISBN 0-89006-861-5, Artech house inc., London.

Appendix 1 : Summary of the Developed Algorithms

The Algorithms can be classified into modules performing different functions or sub-algorithms excluding the Applanix Pos Proc software solution. The modules are summarised in Figure (1) and explained in detail in the following Appendices.

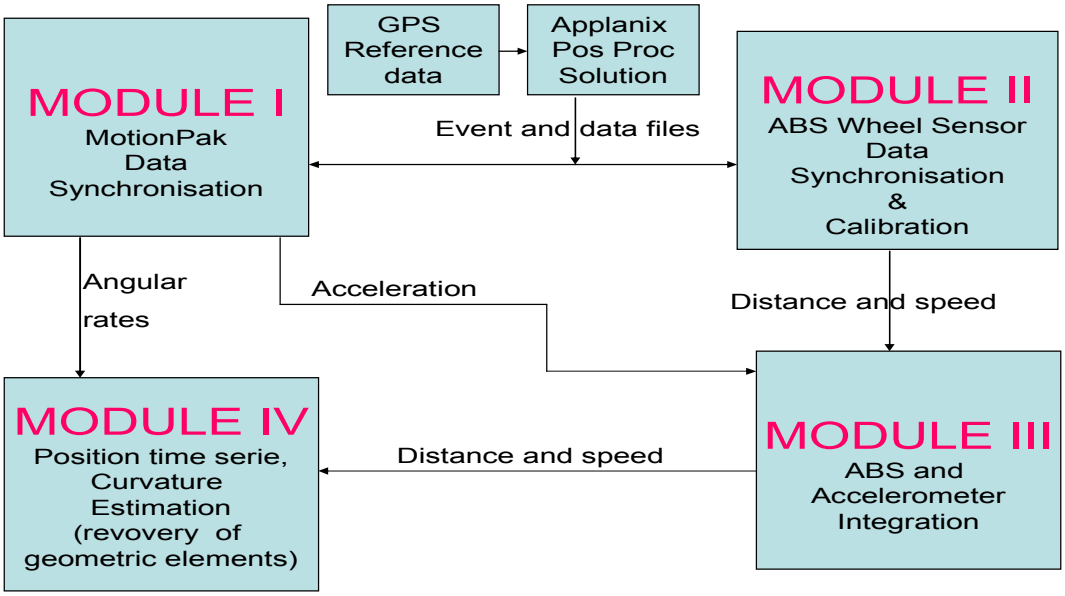

Figure A.1: *The Developed Software Modules*

Module I

The first module is concerned with the processing of the MotionPak and Applanix system measurements and is partly discussed in Section (6.2.2). The Applanix event files are used to compute the time transformation parameters for converting the MotionPak PC time to GPS time. Interpolation is then used to estimate the MotionPak data values at the exact time as the Applanix data times. The synchronized MotionPak measurements are then compensated for pitch angle, gravitation and the bias error. The bias error is estimated during the periods at which the measuring vehicle is at standstill. The *x*-axis accelerations are then forwarded to Module III while the angular rates are passed over to module IV.

Module II

The second module processes the ABS wheel speed sensors' and Applanix system measurements. Using the event file, the ABS measurements are tested and synchronized as described briefly in section (6.1.2). Using the synchronized ABS wheel speed sensors' measurements and the position time series generated by the Applanix *Pos Proc* solution, each of the four ABS wheel speed sensors are calibrated through a loosely coupled Kalman filter derived as detailed in Section (6.2.3). The derived distances and speed are forwarded to Module III.

Module III

The third module integrates the ABS and the single-axis accelerometer measurements using the loosely coupled Kalman filter derived in section (6.3) and listed in Appendix (2). The simulated data was used to test the integration algorithm before being utilised to integrate the field measurements collected by the measuring van around Stuttgart city. The algorithm utilizes the acceleration, single integration of the accelerations, double integration together with the distances and speed estimated by the ABS system. The distances and speed derived from the Integration process are then forwarded to Module IV.

Module IV

The fourth module can be considered as the core module and estimates the position time series of the vehicle track from the solutions of Module I and III. From the position time series, the curvature values along the measured vehicle track are computed using the spline-based functions as detailed in Appendix (3). The estimated curvatures are then analysed to determine the road design elements.

Appendix 2: Kalman Filter Algorithm for the Integration of ABS and Accelerometer

Kalman filter published by Kalman (1960) is a set of mathematical equations that provide an efficient recursive means to estimate the state process by minimizing the squared error. The Kalman filter process estimates the state x of a discrete time controlled process governed by linear stochastic difference equation

$$x_n = Ax_{n-1} + B\varepsilon_n + \omega_n \tag{A2.1}$$

With a measurement given as

$$y_n = Hx_n + v_n \tag{A2.2}$$

The random variables ε_n and ω_n represent the process and measurement noise respectively. They are assumed to be independent of each other, white and with normal probability distributions

$$p(\omega) \sim N(0, Cx_n), \tag{A2.3}$$

$$p(v) \sim N(0, Cv_n). \tag{A2.4}$$

The process noise covariance matrix Cx_n , measurement noise covariance matrix Cv_n, the matrix A relating the state at previous time step n to the state at current time step $n-1$, and the matrix H relating the state to the measurements y_n may all change with each time step. This is the case referred to as extended Kalman filter. The *a priori* and *a posteriori* estimate errors are defined as

$$e_n(-) = x_n - \hat{x}_n(-) \tag{A2.5}$$

$$e_n(+) = x_n - \hat{x}_n(+) \tag{A2.6}$$

The *a priori* estimate error covariance matrix $Cx_n(-)$ is then

$$Cx_n(-) = E\left[(x - \hat{x}_n)(x - \hat{x}_n)^T\right] \tag{A2.7}$$

The *a posteriori* estimate error covariance matrix $Cx_n(+)$ is then

$$Cx_n(+) = E\left[(x - \hat{x}_n)(x - \hat{x}_n)^T\right] \tag{A2.8}$$

A posteriori state estimate $\hat{x}_n$ is then given as

$$\hat{x}_n(+) = \hat{x}_n(-) + K\left(y_n - H\hat{x}_n(-)\right) \tag{A2.9}$$

where the Kalman gain K_n is defined as in Equation (6.45). Then the update equations are

$$\hat{x}_n(-) = A\hat{x}_{n-1} + Bu_{n-1} \tag{A2.10}$$

$$Cx_n(-) = ACx_{n-1}A^T + Cv_n \tag{A2.11}$$

The algorithm that integrates the ABS and accelerometer measurements by using the accelerations to blend the truncated ABS signals is described below.

1. The distance increments and the vehicle speed are derived from the ABS wheel speed sensor measurements through a calibration process described in Section (6.2) i.e. for each wheel sensor $\Delta d = s \cdot \Delta p$ and $\Delta v = s \cdot \Delta p / \Delta t$ where s are the respective scale factors and the Δp signal differences for at every interval Δt.

2. The accelerometer measurements are integrated once to obtain Δd and twice to obtain the velocity difference Δv.
3. Formulation of the linear model listed in Equation (6.37)
4. Initialisation of the Kalman filter parameters
5. A loosely coupled Kalman filter algorithm listed below is used in the integration of the ABS and accelerometer measurements iteratively.

Appendix 3: Spline-Based Curvature Estimation Algorithm

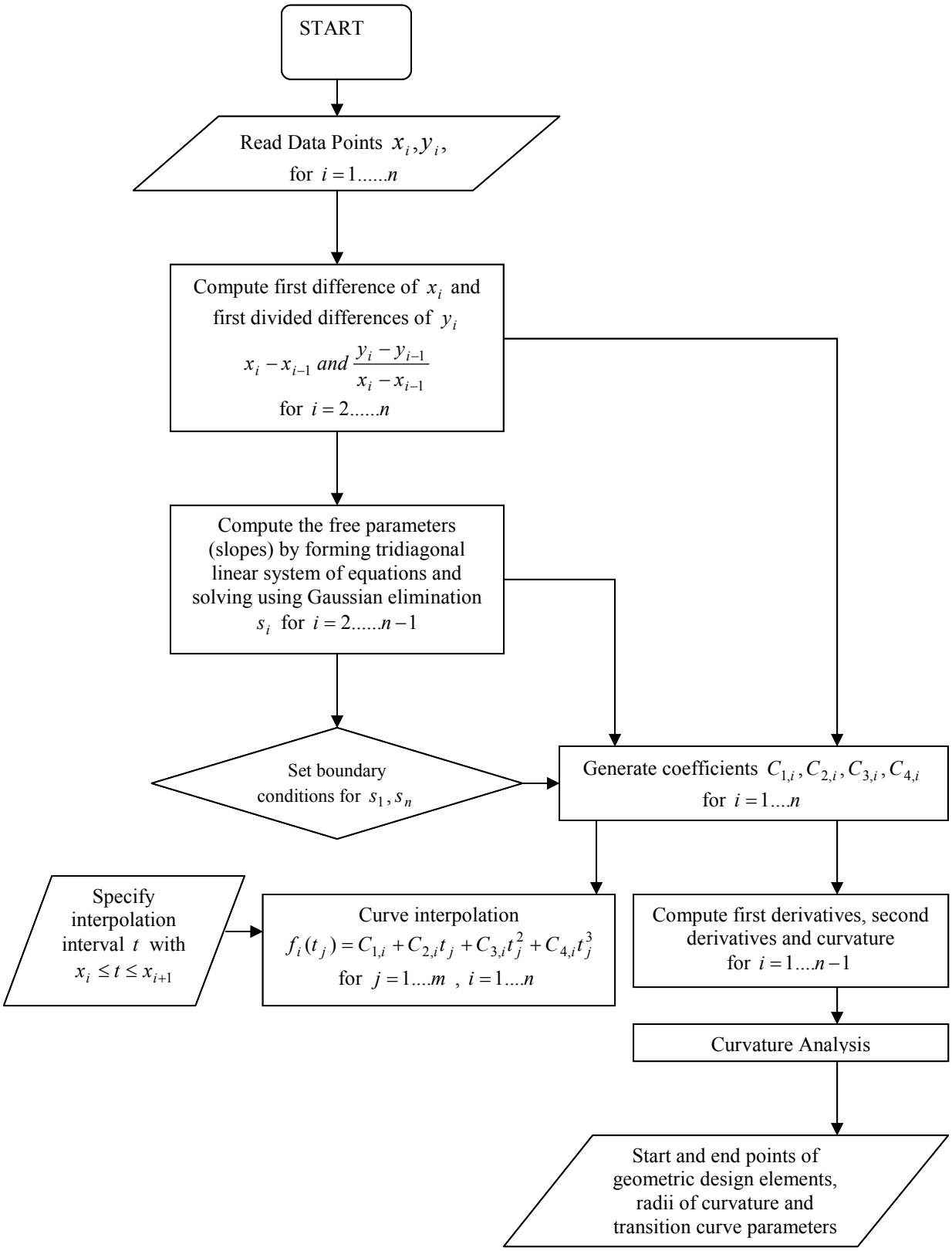

Figure 1*: Curvature computation algorithm using cubic spline interpolation*

1. The first step is to compute the coordinate differences from the measured or simulated x and y coordinates and obtain Δx and Δy
2. The second step involves the curve parameterization using the curve length as briefly stated in Section (4.5). This involves computation of the total curve length and subdividing by the interpolation interval. That is, one computes the distance increments from the coordinate differences, sum them and divide by the chosen spline interpolation interval, say m .

$$\Delta\tilde{d}_i = \sqrt{\left(\Delta\tilde{x}_i^2 + \Delta\tilde{y}_i^2\right)} \Rightarrow D = \sum_{i=1}^{n} \Delta\tilde{d}_i \Rightarrow \Delta d = \frac{D}{m}$$

3. Cubic spline interpolation is then performed using the uniform distant increment Δd as the interpolant for each of the coordinates, by computing differences and the divided differences as shown in Figure (2) above and described in Section (4.3.3.4). That is;

$$c_x = f\left(\Delta d, \Delta\tilde{x}\right) \quad \text{and} \quad c_y = f\left(\Delta d, \Delta\tilde{y}\right)$$

4. Therefore, for each coordinate axis, the divided difference is computed as shown below;

$$\text{div}_{\text{x}} = \frac{\Delta\tilde{x}_i}{\Delta d} \quad \text{and} \quad \text{div}_{\text{y}} = \frac{\Delta\tilde{y}_i}{\Delta d}$$

5. Computation of the cubic spline coefficients as described in Section (4.3.3.4) to obtain the coefficients $c_{x,1}, c_{x,2}, c_{x,3}, c_{x,4}$ and $c_{y,1}, c_{y,2}, c_{y,3}, c_{y,4}$ for the x and y coordinates respectively. The spline fitted coordinate using the knot-a-knot end condition are then estimated;

$$x_i = c_{x,1} + c_{x,2} + c_{x,3} + c_{x,4} \quad \text{and} \quad y_i = c_{y,1} + c_{y,2} + c_{y,3} + c_{y,4}$$

6. The first and second derivative of x'' and y'' coordinates are then computed

$$x_i' = c_{x,2} + 2c_{x,3} + 3c_{x,4} \quad \text{and} \quad y_i' = c_{y,2} + 2c_{y,3} + 3c_{y,4}$$

$$x_i'' = 2c_{x,3} + 6c_{x,4} \quad \text{and} \quad y_i'' = 2c_{y,3} + 6c_{y,4}$$

7. The curvature κ are then computed using Equation (4.62)

$$\kappa = \frac{x'y'' - y'x''}{\left(x'^2 + y'^2\right)^{3/2}}$$